PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 15/86, 5/10, A01K 67/027, C07K 14/72, A61K 48/00, G01N 33/566

A1

(11) International Publication Number: **WO 99/02683**

(43) International Publication Date: 21 January 1999 (21.01.99)

(21) International Application Number: PCT/US98/14215

(22) International Filing Date: 10 July 1998 (10.07.98)

(30) Priority Data:

08/891,298	10 July 1997 (10.07.97)	US
60/091,874	7 July 1998 (07.07.98)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	08/891,298 (CIP)
Filed on	10 July 1997 (10.07.97)
US	60/091,874 (CIP)
Filed on	7 July 1998 (07.07.98)

(71) Applicant *(for all designated States except US)*: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES [US/US]; 10010 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: GAGE, Fred, H. [US/US]; 6950 Avenida Mirola, La Jolla, CA 92037 (US). SUHR, Steven, T. [US/US]; 8360 Via Sonoma #A, La Jolla, CA 92037 (US).

(74) Agent: REITER, Stephen, E.; Gray Cary Ware & Freidenrich LLP, Suite 1600, 4365 Executive Drive, San Diego, CA 92121 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

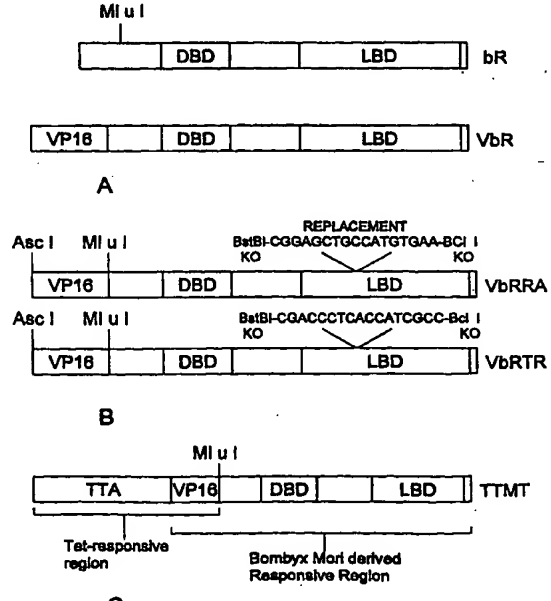
With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MODIFIED LEPIDOPTERAN RECEPTORS AND HYBRID MULTIFUNCTIONAL PROTEINS FOR USE IN REGULATION OF TRANSGENE EXPRESSION

(57) Abstract

In accordance with the present invention, it has been discovered that nuclear receptor proteins isolated from the silk moth *bombyx mori* (bR) are useful for the regulation of transgene expression. bR is generally thought to be a strong transcriptional regulator within cells of the silk moth. In accordance with the present invention, it has been discovered that bR is also functional in mammalian cells. It has further been discovered that the addition of activation domains to the bR open–reading frame enhances the activity of the ligand modulated regulator to afford high–level transcriptional induction. Further modifications to the bR ligand binding domain result in receptors with unique tranactivational characteristics.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

Modified Lepidopteran Receptors and Hybrid Multifunctional Proteins for Use in Regulation of Transgene Expression

ACKNOWLEDGMENT

This invention was made with Government support under Grant No. AG 10435, awarded by the National Institutes of Health. The Government has certain rights in the invention.

RELATED APPLICATIONS

This application claims priority from Provisional Application, U.S.S.N. 60/_____, filed July 7, 1998, and also claims priority from United States Serial No. 08/891,298, filed July 10, 1997, now pending, the entire contents of both of which are hereby incorporated by reference herein.

FIELD OF THE INVENTION

The present invention relates to methods in the field of recombinant DNA technology, and products related thereto. In a particular aspect, the invention relates to methods for modulating the expression of exogenous genes in mammalian systems, and products useful therefor.

BACKGROUND OF THE INVENTION

In the field of genetic engineering, precise control of gene expression is an invaluable tool for studying, manipulating and controlling development and other physiological processes. For example applications for regulated gene expression in mammalian systems include inducible gene targeting, overexpression of toxic and teratogenic genes, anti-sense RNA expression, and gene

therapy (see, for example, Jaenisch, R. (1988) *Science* 240, 1468-1474). For cultured cells, glucocorticoids and other steroids have been used to induce the expression of a desired gene.

As another means for controlling gene expression in mammalian systems, an inducible tetracycline regulated system has been devised and utilized in transgenic mice, whereby gene activity is induced in the absence of the antibiotic and repressed in its presence (see, e.g, Gossen et al. (1992) *PNAS* 89, 5547-5551; Gossen et al.(1993) *TIBS* 18, 471-475; Furth et al. (1994) *PNAS* 91, 9302-9306; and Shockett et al. (1995) *PNAS* 92, 6522-6526). However, disadvantages of the inducible tetracycline system include the requirement for continuous administration of tetracycline to repress expression and the slow clearance of antibiotic from bone, which interferes with regulation of gene expression. While this system has been improved by the recent identification of a mutant tetracycline repressor which acts conversely as an inducible activator, the pharmacokinetics of tetracycline may hinder its use during development when a precise and efficient "on-off" switch is essential (see, e.g., Gossen et al. (1995) *Science* 268, 1766-1769).

Historically, the expression of transgenes with potential pathogenic and cytotoxic properties has been difficult in the context of stable cell lines since cells chronically producing significant levels of the toxic transgene generally die from the cytopathic effects. For example, CAG repeats and concomitant polyglutamine (polyQ) expression has been linked to a variety of neurodegenerative conditions, including Hungtington's disease (Koide et al. (1994) Nat Genet 6:9-13), dentatorubro-pallidoluysian atrophy (Ranum et al. (1994)

Nat Genet 8:280-284), spinocerebellar ataxias (Kawaguch et al. (1994) Nat Genet 8:221-228, Imbert et al. (1993) Nat Genet 4:72-76, Sanpei et al. (1996) Nat Genet 14:277-284, and Zhuchenko et al. (1997) Nat Genet 15:62-69), and spinobulbar muscular atrophy (David et al. (1997) Nat Genet 17:65-70). PolyQ expansion has also been linked to the formation of intracellular protein aggregates *in vivo* in transgenic mice expressing human huntingtin fragments with expanded repeats (Ordway et al (1997) Cell 91:753-763), brain tissue from Huntington's patients (DiFiglia et al. (1997) Science 277:1990-1993) and transiently transfected cultured cells (Martindale et al (1998) Nat Genet 18:150-154 and Cooper et al. (1998) Hum Mol Genet 7:783-790).

In transient transfection models, a 79-Q expanded repeat from the ataxin 3 gene has been shown to form a "punctate pattern" within the cytoplasm, and to result in the apopotic cell death of 70-80% of transfected COS-7 cells within 48 hours of transfection. Paulson et al. has reported that 293 cells transfected with a separate 78-Q ataxin-3 derived construct acquired intranuclear inclusion bodies and also underwent apoptotic cell death in a subset (30%) of the transfected population with 72 hours of transfection. Neuroblastoma cells transiently transfected with an 82-Q construct derived from an expanded huntingtin gene exon 1 were found to be 2-3 times more susceptible to staurosporine induced apoptosis than cells transfected with a LacZ control construct (18), however, spontaneous cell death was not described. Variations in culture techniques, transfection efficiencies, and cell lines, however, make it difficult to rigorously characterize the progression of polyQ expression, intracellular aggregate (IA) formation, and cell death. One impediment to the comparative study of IAs *in vitro* is that transient transfection studies predict that toxicity from polyQ over-expression will prevent the isolation of stable cultured cells.

Accordingly, there is a need in the art for improved systems to precisely modulate the expression of exogenous genes in mammalian subjects. For example, a non-mammalian-based transcription regulating system would be desirable for general application to transgene regulation in *in vitro*, *ex vivo*, and *in vivo* applications, as well as transgenic animals. Such systems would be extremely desirable to conditionally express pathogenic or toxic compounds in order to produce stable cells. A system that is simple, compact and dependent on ligands which are relatively inexpensive, readily available and of low toxicity in animals would prove useful for stimulation of regulated systems.

BRIEF DESCRIPTION OF THE INVENTION

In accordance with the present invention, it has been discovered that nuclear receptor proteins isolated from the silk moth *bombyx mori* (bR) are useful for the regulation of transgene expression. bR is generally thought to be a strong transcriptional regulator within cells of the silk moth. In accordance with the present invention, it has been discovered that bR is also functional in mammalian cells without the addition of exogenous dimer partners therefor. It has further been discovered that the addition of activation domains to the bR open-reading frame (VbR) enhances the activity of the ligand modulated regulator to afford high-level transcriptional induction (see, e.g., Figure 1A). Further modifications to the bR ligand binding domain result in receptors with unique transactivation characteristics (see, e.g., Figure 1B).

In accordance with another aspect of the present invention, hybrid proteins produced by fusion of modified bRs with other ligand-regulated proteins have been found to be capable of high level, regulated transactivation of

sequences controlled by both hormone response elements and tetracycline operators. VbR variants and hybrid proteins (see, e.g., Figure 1C), in combination with the appropriate promoters and transgenes, can be introduced into target cells by common methods such as transfection of plasmids or by virus mediated gene transfer. The small size and simplicity of these proteins makes them particularly attractive for use in retroviral vectors.

Invention receptor peptides have proven capable of providing a high degree (>100 fold) of ligand-dependent transgene expression in a variety of infected and transfected cell types including primary cell types. In addition, several transgenes including luciferase, LacZ, human GH, GFP, tyrosine hydroxylase, and the low-affinity NGF receptor, have all been shown to function in a ligand-dependent fashion in cells infected with invention constructs.

Invention methods provide for regulated gene expression by exogenous non-mammalian inducers, and therefor can be advantageously employed in a variety of *in vivo* and *in vitro* mammalian expression systems. For example, inducible expression of flp recombinase in transgenic mammals, in accordance with invention methods, would enable those of skill in the art to accomplish temporally specific inducible gene targeting of the adult or the developing embryo (See, e.g., U.S. Patent Nos. 5,654,182 and 5,677,177). Stable cell lines conditionally expressing desired genes provide a valuable tool for the study and treatment of diseases. In order to maximize the chance of obtaining stably transfected/infected cultured cells expressing a desired transgene, such as polyglutamine expression cassette (*PQEC*), a retroviral construct is provided in which expression of the transgene is controlled by an exogenous ligand.

For example, in order to study possible mechanisms underlying Huntington's disease, cultured cell types that conditionally express the glutamine (CAG) repeats characteristic of the pathological form of the huntingtin protein are produced. Two aspects of polyQ-mediated cytotoxicity are well-suited to characterization *in vitro;* the expression, distribution, and aggregation of proteins with expanded polyQ tracts within individual cells, and the progression of polyQ expression to death of the cell. The huntingtin protein is expressed predominantly in the cytoplasm of both neuronal and non-neuronal cell types. The expression of polyglutamine tracts separately from the huntingtin protein will allow one to determine what toxic properties, if any, polyglutamine tracts themselves may harbor. A variable length polyglutamine fused to cytosolic reporter proteins such as green fluorescent protein (GFP) facilitates the selection and characterization of *PQEC* ligand-responsive cell populations and clonal lines. The use of recombinant retroviruses as *PQEC* transfer vectors will allow the stable integration of the transgene into a wide variety of cultured cell types.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1A provides a schematic drawing of the nuclear receptor protein isolated from the silk moth *bombyx mori* (bR), as well as a construct containing a VP16 activation domain fused at an internal MluI site near the N-terminus of the *bombyx mori* nuclear receptor, referred to as VbR.

Figure 1B depicts the replacement of several amino acids of the ligand binding domain of the *bombyx mori* nuclear receptor with sequences from the retinoic acid

receptor or the thyroid hormone receptor (VbRRA and VbRTR, respectively).

Figure 1C provides a schematic of a hybrid transactivating construct, designated TTMT, containing a tetracycline-responsive transactivator (Tet-responsive region), a VP16 activation domain and a peptide derived from *bombyx mori*.

Figure 2A shows the induction of fusion proteins (VP16 activation domain fused to either Drosophila ecdysone receptor (dVEcR) or *bombyx mori* nuclear receptor derived(VbR)) with 5 μg/ml tebufenozide in the absence of exogenous dimer partners.

Figure 2B shows that a peptide derived from *bombyx mori* responds to tebufenozide (teb) with greater than 200-fold induction, while Drosophila EcR responds weakly, it at all.

Figure 2C illustrates the action of tebufenozide (teb) on the expression of EcRE-Luc reporter when co-transfected with the following receptors:

LINX: which encodes the tet-transactivator (TTA),
VbR: which encodes VbR,
TTMT: which encodes the TTA-VbR fusion protein, and
TTMT-2V: similar to TTMT but with two VP16 t-domains.

Figure 2D depicts the same experiment illustrated in Figure 2C, utilizing a TetO-luc reporter, which responds only to TTA or fusion proteins in the absence of doxycycline.

Figure 2E shows the results of an experiment employing CV-1 cells transfected with TTMT and two reporters: 6TetO-luc and 4-EcRE-β-galactosidase.

Figure 3 shows that lower levels of teb are as effective as higher levels for stimulating full activity.

Figure 4 is a schematic drawing of representative single-plasmid constructs involving VbR use.

Figure 5 illustrates the construction of the pBW plasmid.

Figure 6A provides a comparison of Bombyx mori ecdysone receptor (BE) and Drosophila melanogaster ecdysone receptor (DE) by transient transfection with both receptors and the dimer partners Usp or RXR in 293 cells.

Figure 6B provides a similar comparison as in Figure 6A, however, in CV-1 cells.

Figure 7A is a schematic drawing of native and chimeric ecdysone receptors.

Figure 7B is a graph depicting the transient transfection of VEH into CV-1 cells.

Figure 8 is a schematic drawing comparing several different chimeric receptors.

Figure 9A provides a schematic of receptor variant viral vectors and an ecdysone response element responsive reporter virus (MS).

Figure 9B illustrates induction levels of luciferase in 293, CV-1 and primary fibroblast cells infected with MS-luc virus.

Figure 10 illustrates the construction of the Boris I plasmid.

Figure 11 illustrates the construction of the Boris III plasmid.

Figure 12 illustrates the construction of the Boris IV plasmid.

Figure 13 illustrates the construction of the pBW-6E-GFP-CB-VBE plasmid.

Figure 14 illustrates the retroviral construct, BORIS III transcribes a transgene in the presence and absence of a ligand.

Figures 15A and B illustrate the construction of a PolyQ expression construct.

Figure 16 provides a comparison of various polyQ-GFP vectors constructed.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, there are provided nucleic acid constructs encoding a receptor peptide comprising:

(i) a ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily,

(ii) a DNA binding domain, and

(iii) an activation domain,

wherein the receptor peptide activates a regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain. Alternatively, there are provided nucleic acid constructs encoding a receptor peptide comprising:

(i) a ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily, and

(ii) a DNA binding domain, wherein the DNA binding domain is not obtained from the same member as the ligand binding domain and hinge region,

wherein the receptor peptide activates a regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

In accordance with another embodiment of the present invention, there are provided method(s) employing such constructs. For example, there are provided method(s) for modulating the transcription of exogenous nucleic acid(s) in a host containing:

(i) a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element; and

(ii) a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer

partner therefor and in the presence of a ligand for said ligand binding domain,

said method comprising administering to said host an amount of ligand effective to modulate the transcription of said exogenous nucleic acid(s); wherein said ligand is not normally present in the cells of said host.

As employed herein, the terms "modulate" and "modulating" refer to the ability of a given ligand/receptor peptide complex to activate/deactivate and/or up-regulate/down-regulate transcription of exogenous nucleic acids, relative to the transactivation activity of said receptor peptide in the absence of ligand. The actual effect of complex formation on the transactivation activity of a receptor peptide will vary depending on the specific ligand and DNA binding domains employed in the receptor peptide and on the regulatory element with which the ligand/receptor peptide complex interacts.

As employed herein, the term "host" refers to the cell, tissue, organ or organism in need of transcriptional regulation of exogenous or endogenous nucleic acids. Preferably, hosts are mammalian or mammalian derived cells or tissue. Exemplary mammals include: humans; domesticated animals, e.g., rat, mouse, rabbit, canine, and feline; farm animals, e.g., chicken, bovine, porcine and ovine; and animals of zoological interest, e.g., monkey and baboon; and the like.

As employed herein, a "dimer partner" refers to members of the nuclear receptor superfamily to which other members preferentially bind to form heterodimeric species. For example, members of the nuclear receptor superfamily preferentially form heterodimers with a common partner, the retinoid X (or 9-cis retinoic acid) receptor (RXR, see, for

example, Yu et al. (1991) *Cell* 67:1251-1266; Bugge et al. (1992) *EMBO J.* 11:1409-1418; Kliewer et al. (1992) *Nature* 355:446-449; Leid et al, (1992) *Cell* 68:377-395; Marks et al. (1992) *EMBO J.* 11:1419-1435; Zhang et al. (1992) *Nature* 355:441-446; Issemann et al. (1993) *Biochimie*. 75:251-256). Additional dimer partners for members of the nuclear receptor superfamily include ultraspiracle (Usp), farnesoid X receptor (FXR), and the like.

Receptor peptides utilized in the present invention are characterized as being fully functional in mammalian cells without the addition of any exogenous dimer partners therefor. For such receptor peptides, the presence of native level or concentration of endogenous dimer partner is sufficient to promote transcription. As employed herein, the phrase "exogenous dimer partner" refers to a dimer partner for the receptor peptide that is introduced into the host. Exogenous dimer partner may be employed because the dimer partner is not present in the host cell (i.e., Usp) or because the dimer partner is required at elevated levels in the host (i.e., higher levels of RXR). Oppositely, the phrase "endogenous dimer partner(s)" refers to dimer partners which are native to the unmodified mammalian host cell. Invention receptors are advantageous because dimer partner is not necessary at elevated levels in order for formation of heterodimers to occur, thereby promoting transcription.

Members of the nuclear receptor superfamily are characterized by the presence of five domains: A/B, C, D, E and F (Evans, R. *Science* 240:889-895 (1988)), wherein "C" corresponds to the DNA binding domain, "D" corresponds to the hinge region, and "E" corresponds to the ligand binding domain. In accordance with a particular aspect of the invention, it has been discovered that the hinge region

(D), in coordination with the ligand binding domain (E), of certain non-mammalian members of the nuclear receptor superfamily enable such members to interact with dimer partners endogenous to mammalian cells, i.e., in the absence of and without requiring introduction of exogenous dimer partners. Specifically, it has been identified that regions within the hinge region ("D") and the ligand binding domain ("E") of the *bombyx mori* receptor function in concert to produce high affinity, ligand-dependent heterodimerization between bR and mammalian dimer partners, without requiring the introduction of exogenous dimer partners such as Usp.

Thus, in accordance with a preferred embodiment of the present invention, the invention receptor peptide comprises the hinge region and ligand binding domain derived from insect receptors of lepidopteran species such as *bombyx mori*, (Swevers et al. *Insect Biochem. Molec. Biol.* 25(7):857-866 (1995)), *Choristoneura fumiferana* (Palli et al. *Insect Biochem. Molec. Biol.* 26(5):485-499 (1996)), *Manduca sexta* (Fujiwara et al. *Insect Biochem. Molec. Biol.* 25(7):845-856 (1995)), *Aedes aegypti* (Cho et al. *Insect Biochem Molec. Biol.* 25:19-27 (1995), *Chorinomus tentans* (Imhof et al. *Insect Biochem. Molec. Biol.* 25:115-124 (1993), and the like. Presently preferred insect receptors from which the hinge region and/or ligand binding domain is derived substantially lack the C-terminal "F" domain. In addition, presently preferred insect receptors share less than 50% sequence similarity with Drosophila ecdysone receptors, comparisons performed by Higgins and Sharp sequence alignment.

As employed herein, the term "hinge region" or "D domain" refers to the domain located between the DNA binding domain and the ligand binding domain of intact

members of the nuclear receptor superfamily. In accordance with a preferred embodiment of the present invention, the hinge region facilitates high affinity interaction with dimer partners endogenous to mammalian cells, also referred herein as "endogenous dimer partners." Preferably, the hinge region of a non-mammalian member of the nuclear receptor superfamily can be characterized as any sequence having substantial sequence identity with amino acid residues 273-362 set forth in SEQ ID NO:3, or substantial portions thereof which confer sufficiently high affinity for endogenous dimer partner.

The hinge region can be functionally located in either orientation and at various positions within the receptor peptide. For example, the hinge region can be positioned at either the amino or carboxy terminus of the receptor peptide, or therebetween. In a preferred embodiment of the present invention, the hinge region is positioned internally between the ligand binding and DNA binding domains of the receptor peptide (see Figure 1A).

As employed herein, the phrase "ligand binding domain of a non-mammalian member of the nuclear receptor superfamily" refers to ligand binding domains derived from receptors which are not endogenous to a mammalian host. Ligand binding domains which are not endogenous to a mammalian host include ligand binding domains which are modified to be non-responsive to ligands endogenous or native to the host. Ligand binding domains contemplated for use according to the present invention can be derived from non-mammalian member(s) of the nuclear receptor superfamily which members are not normally present in the cells of a host. Ligand binding domains are preferably derived from the carboxy-terminal portion of non-mammalian receptors which are capable of activating transcription of regulatory elements in the absence of exogenous dimer

partner therefore. Exemplary receptors which are not normally present in mammalian cells include insect receptors, plant receptors, and the like.

Ligand binding domains can be functionally located in either orientation and at various positions within the receptor peptide. For example, the ligand binding domain can be positioned at either the amino or carboxy terminus of the receptor peptide, or therebetween. In a preferred embodiment of the present invention, the ligand binding domain is positioned at the carboxy terminus of the receptor peptide (see Figure 1A).

Exemplary ligand binding domains can alternatively be characterized as comprising substantial sequence identity with amino acid residues 393-508 ("E_2"), preferably amino acid residues 393-586 (E_2 and E_3, wherein "E_3" is amino acid residues 509-586), as set forth in SEQ ID NO:3, or substantial portions thereof (i.e., typically at least 46 or more contiguous nucleotides thereof). It has been determined that the specific ligand response determinant of bombyx lies within the E_2 region of the native bR, whereas both E_2 and E_3 regions facilitate heterodimer formation with endogenous dimer partners. Modifications of this sequence contemplated for use in the practice of the present invention include replacing several amino acids of the ligand binding domain with sequences from ligand binding domains of other members of the nuclear receptor superfamily, such as retinoic acid receptor and/or thyroid hormone receptor (Figure 1B). These modifications provide unique transactivating characteristics and/or eliminate restriction sites, which facilitate the construction of useful peptide-based viral constructs.

DNA-binding domains contemplated for use in the preparation of invention receptor peptides are typically

obtained from DNA-binding proteins (e.g., transcription factors). The term "DNA-binding domain" is understood in the art to refer to an amino acid sequence that is able to bind to DNA. As used herein, the term "DNA-binding domain" encompasses a minimal peptide sequence of a DNA-binding protein up to the entire length of a DNA-binding protein, so long as the DNA-binding domain functions to associate with a particular regulatory element.

DNA-binding domains are known to function heterologously in combination with other functional domains by maintaining the ability to bind DNA recognition sequence (see, e.g., Brent and Ptashne, (1985) *Cell*, 43:729-736). For example, with respect to hormone receptors, DNA-binding domains are interchangeable, thereby providing numerous chimeric receptor proteins (see, e.g., U.S. Patent 4,981,784; and Evans, R., (1988) *Science*, 240:889-895). Similar to the ligand binding domain, the DNA-binding domain can be positioned at either the carboxy terminus or the amino terminus, or the DNA-binding domain can be positioned between the ligand binding domain and the activation domain. In preferred embodiments of the present invention, the DNA-binding domain is positioned internally between the ligand binding domain and the activation domain.

"DNA-binding protein(s)" contemplated for use herein belong to the well-known class of proteins that are able to directly bind DNA and facilitate initiation or repression of transcription. Exemplary DNA-binding proteins contemplated for use herein include transcription control proteins (e.g., transcription factors and the like; Conaway and Conaway, 1994, "Transcription Mechanisms and Regulation", *Raven Press Series on Molecular and Cellular Biology*, Vol. 3, Raven Press, Ltd., New York, NY).

Transcription factors contemplated for use herein as a source of such DNA binding domains include, e.g., homeobox proteins, zinc finger proteins, hormone receptors, helix-turn-helix proteins, helix-loop-helix proteins, basic-Zip proteins (bZip), β-ribbon factors, and the like. See, for example, Harrison, S., "A Structural Taxonomy of DNA-binding Domains," *Nature*, 353:715-719. Homeobox DNA-binding proteins suitable for use herein include, for example, HOX, STF-1 (Leonard et al., 1993, *Mol. Endo.*, 7:1275-1283), Antp, Mat α-2, INV, and the like. See, also, Scott et al. (1989), *Biochem. Biophys. Acta*, 989:25-48. It has been found that a fragment of 76 amino acids (corresponding to amino acids 140-215 described in Leonard et al., (1993) *Mol. Endo.*, 7:1275-1283) containing the STF-1 homeodomain binds DNA as tightly as wild-type STF-1. Suitable zinc finger DNA-binding proteins for use herein include Zif268, GLI, XFin, and the like. See also, Klug and Rhodes (1987) *Trends Biochem. Sci.*, 12:464; Jacobs and Michaels (1990) *New Biol.*, 2:583; and Jacobs (1992), *EMBO J.*, 11:4507-4517.

An additional DNA binding domain contemplated for use in the practice of the present invention is the GAL4 DNA binding domain. The DNA binding domain of the yeast GAL4 protein comprises at least the first 74 amino terminal amino acids thereof (see, for example, Keegan et al., *Science* 231:699-704 (1986)). Preferably, the first 90 or more amino terminal amino acids of the GAL4 protein will be used, with the 147 amino terminal amino acid residues of yeast GAL4 being presently most preferred.

Preferably, DNA-binding domain(s) used herein is(are) obtained from a member of the nuclear receptor superfamily. As used herein, the phrase "member(s) of the nuclear receptor superfamily" (also known as "intracellular

receptors" or "steroid/thyroid hormone superfamily of receptors") refers to hormone binding proteins that operate as ligand-dependent transcription factors, including identified members of the nuclear receptor superfamily for which specific ligands have not yet been identified (referred to hereinafter as "orphan receptors").

Exemplary members of the steroid/thyroid hormone superfamily of receptors (including the various isoforms thereof) include steroid receptors such as glucocorticoid receptor (GR), mineralocorticoid receptor (MR), estrogen receptor (ER), progesterone receptor (PR), androgen receptor (AR), vitamin D_3 receptor (VDR), various isoforms of peroxisome proliferator-activated receptors (PPARs), and the like; plus retinoid receptors, such as the various isoforms of retinoic acid receptor (e.g., RARα, RARβ, or RARγ), the various isoforms of retinoid X receptor (e.g., RXRα, RXRβ, or RXRγ), and the like (see, e.g., U.S. Patent Nos. 4,981,784; 5,171,671; and 5,071,773); thyroid receptors (TR), such as TRα, TRβ, and the like; insect derived receptors such as the ecdysone receptor, and the like; as well as other gene products which, by their structure and properties, are considered to be members of the superfamily, as defined hereinabove, including the various isoforms thereof. Examples of orphan receptors contemplated for use herein as a source of DNA binding domain include HNF4 (see, for example, Sladek et al., in *Genes & Development* **4**: 2353-2365 (1990)), the COUP family of receptors (see, for example, Miyajima et al., in *Nucleic Acids Research* **16**: 11057-11074 (1988), and Wang et al., in *Nature* **340**: 163-166 (1989)), COUP-like receptors and COUP homologs, such as those described by Mlodzik et al., in *Cell* **60**: 211-224 (1990) and Ladias et al., in *Science* **251**: 561-565 (1991), the insect derived knirps and knirps-related receptors, and the like.

The DNA-binding domains of all members of the nuclear receptor superfamily are related. Such domains consist of 66-68 amino acid residues, and possess about 20 invariant amino acid residues, including nine cysteines. Members of the superfamily are characterized as proteins which contain these 20 invariant amino acid residues. The highly conserved amino acids of the DNA-binding domain of members of the superfamily are as follows:

Cys - X - X - Cys - X - X - Asp* - X - Ala* - X - Gly* - X - Tyr* - X - X - X - X - Cys - X - X - Cys - Lys* - X - Phe - Phe - X - Arg* - X - X - X - X - X - X - X - X - X - (X - X -) Cys - X -X - X - X - X - (X - X - X -) Cys - X -X - X - Lys - X - X - Arg - X - X - Cys - X - X - Cys - Arg* - X - X - Lys* - Cys - X - X - X - Gly* - Met (SEQ ID NO:1);

wherein X designates non-conserved amino acids within the DNA-binding domain; an asterisk denotes the amino acid residues which are almost universally conserved, but for which variations have been found in some identified hormone receptors; and the residues enclosed in parenthesis are optional residues (thus, the DNA-binding domain is a minimum of 66 amino acids in length, but can contain several additional residues).

It has also been found that *in vitro* evolution methods can be applied to modify and improve existing DNA-binding domains (see, e.g., Devlin et al., 1990, Science, 249:404-406; and Scott and Smith, 1990, Science, 249:386-390). Alternatively, DNA-binding domains which are engineered with novel DNA-recognition specificity (see, e.g., Pomerantz et al. Science 267:93-96, 1995, ZFHD1, an engineered transcription factor with a composite DNA-binding domain) are also contemplated.

Two polynucleotides or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues in the two sequences is the same when aligned for maximum correspondence. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, *Adv. Appl. Math.*, **2:482 (1981)**, by the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.*, **48:443 (1970)**, by the search for similarity method of Pearson and Lipman, *Proc. Natl. Acad. Sci.* (U.S.A.), **85:2444 (1988)**, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by inspection. These references are incorporated herein by reference.

The percentage of sequence identity between two sequences is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity.

For instance, a preferred method for comparing sequences uses the GAP program based on the algorithm of Needleman, et at., supra. Typically, the default values for all parameters are selected. These are gap weight: 5.0, length weight: 0.30, average match: 1.0, and average mismatch: 0.0.

As used herein, the phrase "substantial sequence identity" refers to nucleotide or amino acid sequences which share at least 80% sequence identity, preferably 90%, more preferably 95 % or more, regardless of the algorithm used to determine sequence identity, compared to a reference sequence over a comparison window of about 20 bp to about 2000 bp, typically about 50 to about 1500 bp, usually about 350 bp to about 1200. The values of percent identity are preferably determined using the GAP program, referred to above. Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. It is recognized, however, that proteins (and DNA or mRNA encoding such proteins) containing less than the above-described level of homology produced as splice variants or as a result of conservative amino acid substitutions (or substitution of degenerate codons) are contemplated to be within the scope of the present invention.

The phrase "substantially the same," used in reference to a DNA nucleotide sequence, an RNA ribonucleotide sequence, or an amino acid sequence, refers to sequences that have slight and non-consequential sequence variations from the actual sequences disclosed herein. Species that are substantially the same are considered to be equivalent to the disclosed sequences. In this regard, "slight and non-consequential sequence variations" mean that sequences that are substantially the same as the DNA, RNA, or proteins disclosed and claimed herein are functionally equivalent to the sequences disclosed and claimed herein. Functionally equivalent sequences will function in substantially the same manner to produce substantially the same effect as the sequences disclosed.

Receptor peptides employed in the present invention can be modified by the introduction of activation domains.

Activation domains contemplated for use in the practice of the present invention are well known in the art and can readily be identified by those of skill in the art. Activation domains contemplated for use herein are typically derived from transcription factors and comprise a contiguous sequence that functions to activate gene expression when associated with a suitable DNA-binding domain and a suitable ligand binding domain. An activation domain can be positioned at any convenient site within the receptor peptide, i.e., at the carboxy terminus, the amino terminus or between the ligand binding domain and the DNA binding domain. In preferred embodiments of the present invention, the activation domain is positioned at the amino terminus of the receptor peptide.

Suitable activation domains can be obtained from a variety of sources, e.g., from the N-terminal region of a member of the steroid/thyroid hormone superfamily of receptors, from a transcription factor activation domain, such as, for example, VP16, GAL4, NF-κB or BP64 activation domains, and the like. The presently most preferred activation domain contemplated for use in the practice of the present invention is obtained from the N-terminal region of the VP16 protein.

In a particular embodiment of the present invention, the invention nucleic acid construct further comprises promoters and regulatory elements operatively associated with exogenous nucleic acids. In a preferred embodiment of the present invention, receptor peptide, in the presence of a ligand therefor, binds the regulatory element and activates transcription of the exogenous nucleic acids.

Regulatory elements contemplated for use in the practice of the present invention include elements responsive to the invention receptor peptide. In a

preferred embodiment of the present invention, such elements are exogenous regulatory elements not normally present in the cells of the host. One class of exogenous regulatory elements contemplated for use herein includes hormone response elements which modulate transcription of exogenous nucleic acid when bound to the DNA binding domain of an invention receptor peptide.

Additional regulatory elements that may be utilized in the practice of the present invention include exogenous regulatory elements responsive to a non-mammalian transactivator. One such transactivator-responsive regulatory element is an operator which confers responsiveness to antibiotics. Exemplary operators contemplated for use in this aspect of the invention include the tetracycline-analog regulated operator, the TET operator, the Lac operator, and the like.

Exogenous response elements contemplated for use herein are short cis-acting sequences (i.e., having about 12-20 bp) that are required for activation of transcription in response to association with a suitable ligand, such as diacyl hydrazines, and an invention receptor peptide. Response element sequences contemplated for use herein function in a position- and orientation-independent fashion. Exemplary response elements include hormone response elements, GAL4 response elements and the like.

Hormone response elements contemplated for use in the practice of the present invention are response elements which are responsive to members of the nuclear receptor superfamily. These response elements comprise at least two half-sites (in either direct repeat or inverted repeat orientation to one another), separated by a spacer of 0-5 nucleotides. As used herein, the term "half-site" refers to a contiguous 6 nucleotide sequence that is bound by a

particular member of the nuclear receptor superfamily. Typically, two half-sites with a corresponding spacer make up a hormone response element. Hormone response elements can be incorporated in multiple copies into various transcription regulatory regions.

Preferred hormone response elements employed in the practice of the present invention comprise a first half-site and a second half-site, separated by a spacer of 0-5 nucleotides;

wherein each half-site has the sequence:

-RGBNNM-,

(or complements thereof) wherein

each R is independently selected from A or G;

each B is independently selected from G, C, or T;

each N is independently selected from A, T, C, or G; and

each M is independently selected from A or C;

with the proviso that at least 4 nucleotides of each -RGBNNM- group of nucleotides are identical with the nucleotides at comparable positions of the sequence -AGGTCA-.

Exemplary half-sites having the -RGBNNM- motif for use in preparing response elements useful in the practice of the present invention include, for example, half-sites selected from -AGGGCA-, -AGTTCA-, -AGGTAA-, -AGGTCA-, -GGTTCA-, -GGGTTA-, -GGGTGA-, -AGGTGA-, -GGGTCA-, and the like. A particularly preferred first half-site is -AGTGCA-.

Additional response elements included the GAL4 response element. Exemplary GAL4 response elements are those containing the palindromic 17-mer:

5'-CGGAGGACTGTCCTCCG-3' (SEQ ID NO:4),

such as, for example, 17MX, as described by Webster et al., in Cell 52:169-178 (1988), as well as derivatives thereof. Additional examples of suitable response elements include those described by Hollenberg and Evans in *Cell* 55:899-906 (1988); or Webster et al. in *Cell* 54:199-207 (1988).

Regulatory elements employed in the practice of the present invention are operably linked to a suitable promoter for transcription of exogenous nucleic acid(s) product(s). When exogenous nucleic acid(s), operatively linked to a suitable promoter, is(are) introduced into the cells of a suitable host, expression of the exogenous nucleic acid(s) is(are) controlled by the presence of ligands, which are not normally present in the host cells.

As used herein, when referring to nucleic acids, the phrase "exogenous to said mammalian host" or simply "exogenous" refers to nucleic acids not naturally found at levels sufficient to provide a function in the particular cell where transcription is desired. For example, exogenous nucleic acids can be either natural or synthetic nucleic acids, which are introduced into the host in the form of DNA or RNA. The nucleic acids of interest can be introduced into target cells (for *in vitro* applications), or the nucleic acids of interest can be introduced directly or indirectly into a host by the transfer of transformed cells into a host.

In contrast to exogenous nucleic acids, the phrase "endogenous nucleic acids" or "endogenous genes" refers to nucleic acids naturally found at levels sufficient to provide a function in the particular cell where transcription is desired.

Exogenous nucleic acids contemplated for use in the practice of the present invention include wild type and/or therapeutic nucleic acids.

"Wild type" genes are those that are native to cells of a particular type. Exemplary wild type nucleic acids are genes which encode products:

the substantial absence of which leads to the occurrence of a non-normal state in a host; or

a substantial excess of which leads to the occurrence of a non-normal state in a host.

Such genes may not be expressed in biologically significant levels or may be undesirably overexpressed, respectively. Thus, for example, while a synthetic or natural gene coding for human insulin would be exogenous genetic material to a yeast cell (since yeast cells do not naturally contain insulin genes), a human insulin gene inserted into a human skin fibroblast cell would be a wild type gene with respect to the fibroblast since human skin fibroblasts contain genetic material encoding human insulin, although human skin fibroblasts do not express human insulin in biologically significant levels.

Therapeutic nucleic acids contemplated for use in the practice of the present invention include those which:

encode products which are toxic to the cells in which they are expressed; or

encode products which impart a beneficial property to a host; or

those which transcribe nucleic acids which modulate transcription and/or translation of endogenous genes.

As employed herein, the phrase "therapeutic nucleic acids" refers to nucleic acids which impart a beneficial function to the host in which such nucleic acids are transcribed. Therapeutic nucleic acids are those that are

not naturally found in host cells. For example, synthetic or natural nucleic acids coding for wild type human insulin would be therapeutic when inserted into a skin fibroblast cell so as to be expressed in a human host, where the human host is not otherwise capable of expressing functionally active human insulin in biologically significant levels. Further examples of therapeutic nucleic acids include nucleic acids which transcribe antisense constructs used to suppress the expression of endogenous genes. Such antisense transcripts bind endogenous nucleic acid (mRNA or DNA) and effectively cancel out the expression of the gene. In accordance with the methods described herein, therapeutic nucleic acids are expressed at a level that provides a therapeutically effective amount of the corresponding therapeutic protein.

Exogenous nucleic acids useful in the practice of the present invention include genes that encode biologically active proteins of interest, such as, e.g., secretory proteins that can be released from said cell; enzymes that can metabolize a toxic substance to produce a non-toxic substance, or that metabolize an inactive substance to produce a useful substance; regulatory proteins; cell surface receptors; and the like. Useful genes include genes that encode blood clotting factors such as human factors VIII and IX; genes that encode hormones such as insulin, parathyroid hormone, luteinizing hormone releasing factor (LHRH), alpha and beta seminal inhibins, and human growth hormone; genes that encode proteins such as enzymes, the absence of which leads to the occurrence of an abnormal state; genes encoding cytokines or lymphokines such as interferons, granulocytic macrophage colony stimulating factor (GM-CSF), colony stimulating factor-1 (CSF-1), tumor necrosis factor (TNF), and erythropoietin (EPO); genes encoding inhibitor substances such as $alpha_1$-antitrypsin; genes or nucleotides which characterize specific conditions

or diseases (i.e., CAG repeat expansion and concomitant polyglutamine (polyQ) expression has been linked to a variety of neurodegenerative conditions including Huntington's disease, dentatorubro-pallidoluysian atrophy, spinocerebellar ataxias, and spinobulbar muscular atrophy, and the like); genes encoding substances that function as drugs, e.g., genes encoding the diphtheria and cholera toxins; and the like.

Additional nucleic acids contemplated for use in accordance with the present invention include genes which encode proteins present in dopaminergic neurons (useful, for example, for the treatment of Parkinson's disease), cholinergic neurons (useful, for example, for the treatment of Alzheimer's disease), hippocampal pyramidal neurons (also useful for the treatment of Alzheimer's disease), norepinephrine neurons (useful, for example, for the treatment of epilepsy), spinal neurons (useful, for example, for the treatment of spinal injury), glutamatergic neurons (useful, for example, for the treatment of schizophrenia), cortical neurons (useful, for example, for the treatment of stroke and brain injury), motor and sensory neurons (useful, for example, for the treatment of amyotrophic lateral sclerosis), and the like.

Typically, nucleic acid sequence information for proteins encoded by exogenous nucleic acid(s) contemplated for use employed herein can be located in one of many public access databases, e.g., GENBANK, EMBL, Swiss-Prot, and PIR, or in related journal publications. Thus, those of skill in the art have access to sequence information for virtually all known genes. Those of skill in the art can either obtain the corresponding nucleic acid molecule directly from a public depository or the institution that published the sequence. Optionally, once the nucleic acid sequence encoding a desired protein has been ascertained,

the skilled artisan can employ routine methods, e.g., polymerase chain reaction (PCR) amplification, to isolate the desired nucleic acid molecule from the appropriate nucleic acid library. Thus, all known nucleic acids encoding proteins of interest are available for use in the methods and products described herein.

Additional components which can optionally be incorporated into invention constructs include selectable markers. Selectable markers contemplated for use in the practice of the present invention include radiolabeled molecules, fluorescent molecules, ligands, enzymes, and the like. Preferable selectable markers are enzymes such as antibiotic resistance genes, genes which enable cells to process metabolic intermediaries, and the like. Exemplary antibiotic resistance genes include genes which impart tetracycline resistance, genes which impart ampicillin resistance, zeomycin resistance, neomycin resistance, hygromycin resistance, puromycin resistance, and the like. Genes which enable cells to process metabolic intermediaries include genes which permit cells to incorporate L-histidinol, genes encoding thymidine kinase, genes encoding xanthine-guanine phosphoribosyl transferase (gpt), genes encoding dihydrofolate reductase, genes encoding asparagine synthetase, and the like.

In accordance with a preferred embodiment of the present invention, the invention nucleic acid construct further comprises, in addition to the receptor peptide, a non-mammalian transactivator, not a member of the nuclear receptor superfamily and not normally present in the cells of said host, and a compatible transactivator responsive regulatory element not normally present in cells of said host. The transactivator responsive regulatory element controls transcription of the exogenous nucleic acid(s) or a second nucleic acid construct comprising a second

exogenous nucleic acid(s). Examples of transactivator responsive regulatory elements include operators which are responsive to non-mammalian transactivators which confer responsiveness to antibiotics. Exemplary operators contemplated for use in this aspect of the invention include the tetracycline-analog regulated operator, the TET operator, the Lac operator, and the like. Preferably, the transactivator responsive regulatory elements employed in the practice of the present invention are operably linked to a suitable promoter for transcription of exogenous nucleic acid(s) proteins.

Non-mammalian transactivators, other than members of the nuclear receptor superfamily, contemplated for use in the practice of the present invention function in the absence of exogenous dimer partner. Examples of transactivators that typically function in the absence of exogenous dimer partner are tetracycline-controlled transactivators, Vp16-Lac fusion transactivators, and the like. When contained as part of a transactivating construct, the transactivator can be positioned at any convenient site within the construct, i.e., at the carboxy terminus or the amino terminus of the transactivating construct. In preferred embodiments of present invention, the transactivator is positioned at the amino terminus of the transactivating construct (Figure 1C).

Preferably, the non-mammalian transactivator confers responsiveness to antibiotics. An example of a non-mammalian transactivator which confers responsiveness to antibiotics, as contemplated for use in the practice of the present invention is the tetracycline-controlled transactivator. The tetracycline inducible system is well-known in the art (see, e.g, Gossen et al. (1992) *PNAS* 89, 5547-5551; Gossen et al. (1993) *TIBS* 18, 471-475; Furth et al. (1994) *PNAS* 91, 9302-9306; Shockett et al. (1995) *PNAS*

92, 6522-6526; and Hoshimaru et al. (1996) *PNAS* 93(4):1518-1523). Other examples of non-mammalian transactivators well known in the art include the IPTG inducible system based on a VP16-Lac repressor fusion which functions through lac operator sequences inserted into heterologous promoters (see, e.g., Baim et al. (1991) *PNAS* 88:5072-5076).

In addition to the receptors/transactivators set forth above, those of skill in the art recognize that other transactivators can be used herein, e.g., homeobox proteins, zinc finger proteins, hormone receptors, helix-turn-helix proteins, helix-loop-helix proteins, basic-Zip proteins (bZip), β-ribbon factors, and the like. See, for example, Harrison, S., "A Structural Taxonomy of DNA-binding Domains," *Nature*, 353:715-719. Homeobox DNA-binding proteins suitable for use herein include, for example, HOX, STF-1 (Leonard et al., (1993) *Mol. Endo.*, 7:1275-1283), Antp, Mat α-2, INV, and the like. See, also, Scott et al. (1989) *Biochem. Biophys. Acta*, 989:25-48. It has been found that a fragment of 76 amino acids (corresponding to amino acids 140-215 described in Leonard et al., (1993) *Mol. Endo.*, 7:1275-1283) containing the STF-1 homeodomain binds DNA as tightly as wild-type STF-1. Suitable zinc finger DNA-binding proteins for use herein include Zif268, GLI, XFin, and the like. See also, Klug and Rhodes (1987) *Trends Biochem. Sci.*, 12:464; Jacobs and Michaels (1990) *New Biol.*, 2:583; and Jacobs (1992) *EMBO J.*, 11:4507-4517.

In a preferred embodiment of the present invention, expression cassettes are prepared by operably linking invention nucleic acid constructs to a suitable promoter for expression of the encoded receptor peptide. As used herein, the term "promoter" refers to a specific nucleotide

sequence recognized by RNA polymerase, the enzyme that initiates RNA synthesis. The promoter sequence is the site at which transcription can be specifically initiated under proper conditions. When nucleic acid constructs, operatively linked to a suitable promoter, is introduced into the cells of a suitable host, expression of the nucleic acid construct is conditionally or perpetually initiated.

Promoters contemplated for use in the practice of the present invention include inducible (e.g., minimal CMV promoter, minimal TK promoter, modified MMLV LTR), constitutive (e.g., chicken b-actin promoter, MMLV LTR (non-modified), DHFR), and/or tissue specific promoters.

Inducible promoters contemplated for use in the practice of the present invention comprise transcription regulatory regions that function maximally to promote transcription of mRNA under inducing conditions. Examples of suitable inducible promoters include DNA sequences corresponding to: the *E. coli* lac operator responsive to IPTG (see Nakamura et al., *Cell*, 18:1109-1117, 1979); the metallothionein promoter metal-regulatory-elements responsive to heavy-metal (e.g., zinc) induction (see Evans et al., U.S. Patent No. 4,870,009), the phage T7lac promoter responsive to IPTG (see Studier et al., *Meth. Enzymol.*, 185: 60-89, 1990; and U.S. #4,952,496), the heat-shock promoter; the TK minimal promoter; the CMV minimal promoter; a synthetic promoter; and the like.

Exemplary constitutive promoters contemplated for use in the practice of the present invention include the CMV promoter, the SV40 promoter, the DHFR promoter, the mouse mammary tumor virus (MMTV) steroid-inducible promoter, Moloney murine leukemia virus (MMLV) promoter, elongation

factor 1α (EF1α) promoter, albumin promoter, APO A1 promoter, cyclic AMP dependent kinase II (CaMKII) promoter, keratin promoter, CD3 promoter, immunoglobulin light or heavy chain promoters, neurofiliment promoter, neuron specific enolase promoter, L7 promoter, CD2 promoter, myosin light chain kinase promoter, HOX gene promoter, thymidine kinase (TK) promoter, RNA Pol II promoter, MYOD promoter, MYF5 promoter, phophoglycerokinase (PGK) promoter, Stf1 promoter, Low Density Lipoprotein (LDL) promoter, chicken b-actin promoter (used in conjunction with ecdysone response element) and the like.

As readily understood by those of skill in the art, the term "tissue specific" refers to the substantially exclusive initiation of transcription in the tissue from which a particular promoter, which drives expression of a given gene, is derived (e.g., expressed only in T-cells, endothelial cells, smooth muscle cells, and the like). Exemplary tissue specific promoters contemplated for use in the practice of the present invention include the GH promoter, the NSE promoter, the GFAP promoter, neurotransmitter promoters (e.g., tyrosine hydroxylase, TH, choline acetyltransferase, ChAT, and the like), promoters for neurotropic factors (e.g., a nerve growth factor promoter, NT-3, BDNF promoters, and the like), and so on.

In yet another embodiment of the present invention, there are provided constructs comprising a promoter, a tetracycline-controlled transactivator, a VP16 activation domain, a DNA binding domain and *bombyx mori*-derived hinge region and ligand binding domain encoding sequence, wherein the components of the construct are operatively associated with the other components of the construct.

As used herein, the phrase "operatively associated with" refers to the functional relationship of DNA with

regulatory and effector sequences of nucleotides, such as promoters, enhancers, transcriptional and translational stop sites, and other signal sequences. For example, operative linkage of DNA to a promoter refers to the physical and functional relationship between the DNA and promoter such that the transcription of such DNA is initiated from the promoter by an RNA polymerase that specifically recognizes, binds to and transcribes the DNA.

In yet another embodiment of the present invention, there are provided constructs comprising nucleic acid(s) encoding a a VP16 activation domain operatively associated with nucleic acid encoding a DNA binding domain and *bombyx mori*-derived hinge region and ligand binding domain.

In accordance with another embodiment of the present invention, there are provided gene transfer vectors useful for the introduction of invention constructs into suitable host cells. Such gene transfer vectors preferably comprise a first reporter under the control of a regulatory element, a second reporter under the control of an operator which is responsive to a ligand-mediated receptor which confers responsiveness to antibiotics, and a construct comprising a promoter and nucleic acid encoding a VP16 activation domain, a DNA binding domain and *bombyx mori*-derived hinge region and ligand binding domain, optionally, a tetracycline-controlled transactivator. The number of copies of regulatory elements can readily be varied by those of skill in the art. For example, transcription regulatory regions can contain from 1 up to about 50 copies of a particular regulatory element, preferably 2 up to about 25 copies, more preferably 3 up to about 10-15 copies, with about 4-6 copies being especially preferred.

Gene transfer vectors (also referred to as "expression vectors") contemplated for use herein are recombinant

nucleic acid molecules that are used to transport invention nucleic acid constructs into host cells for expression and/or replication thereof. Expression vectors may be either circular or linear, and are capable of incorporating a variety of nucleic acid constructs therein. Expression vectors typically come in the form of a plasmid that, upon introduction into an appropriate host cell, results in expression of the inserted nucleic acid.

Suitable expression vectors for use herein include a recombinant DNA or RNA construct(s), such as plasmids, phage, recombinant virus or other vectors that, upon introduction into an appropriate host cell, result(s) in expression of the inserted DNA. Appropriate expression vectors are well known to those of skill in the art and include those that are replicable in eukaryotic cells and/or prokaryotic cells and those that remain episomal or those which integrate into the host cell genome.

As used herein, the phrase "transcription regulatory region" refers to that portion of a nucleic acid or gene construct that controls the initiation of mRNA transcription. Regulatory regions contemplated for use herein, in the absence of the non-mammalian transactivator, typically comprise at least a minimal promoter in combination with a regulatory element responsive to the ligand/receptor peptide complex. A minimal promoter, when combined with a regulatory element functions to initiate mRNA transcription in response to a ligand/receptor peptide complex. However, transcription typicallywill not occur unless the required inducer (ligand therefor) is present. Conversely, when the non-mammalian transactivator, other than members of the nuclear receptor superfamily, is present in the host, the transactivator-responsive regulatory element will typically only induce transcription in the absence of ligand therefor.

Preferably, the transcription regulatory region further comprises a binding site for ubiquitous transcription factor(s). Such binding sites are preferably positioned between the promoter and the regulatory element. Suitable ubiquitous transcription factors for use herein are well-known in the art and include, for example, Sp1.

Expression vectors suitable for use in the practice of the present invention are well known to those of skill in the art and include those that are replicable in eukaryotic cells and/or prokaryotic cells as well as those that remain episomal and those that integrate into the host cell genome. Expression vectors typically further contain other functionally important nucleic acid sequences encoding antibiotic resistance proteins, and the like.

Exemplary eukaryotic expression vectors include eukaryotic constructs, such as the pSV-2 gpt system (Mulligan et al., (1979) *Nature*, 277:108-114); pBlueSkript (Stratagene, La Jolla, CA), the expression cloning vector described by Genetics Institute (*Science*, (1985) 228:810-815), and the like. Each of these plasmid vectors are capable of promoting expression of the receptor peptide of interest.

In a specific embodiment, a gene transfer vector contemplated for use herein is a viral vector, such as Adenovirus, adeno-associated virus, or herpes-simplex virus based vectors, and synthetic vectors for gene therapy, and the like (see, e.g., Suhr et al. (1993) *Arch. of Neurol.* 50:1252-1268). Adenovirus and adeno-associated virus are extremely suitable as gene transfer vectors to produce receptor peptides. Alternatively, a gene transfer vector employed herein is a retroviral vector. Retroviral vectors are gene transfer plasmids that have an expression construct containing an exogenous nucleic acid residing

between two retroviral LTRs. Retroviral vectors typically contain appropriate packaging signals (e.g., a retroviral psi (Ψ) packaging signal), elements which enable invention construct integration into a host cell (e.g., a 5' and/or a 3' long terminal repeat (LTR)), and/or genes encoding proteins required for retroviral packaging (e.g., the *pol* gene, the *gag* gene and the *env* gene), that enable the retroviral vector, RNA transcribed using the retroviral vector as a template, to be packaged into a viral virion in an appropriate packaging cell line (see, e.g., U.S. Patent 4,650,764).

Suitable retroviral vectors for use herein are described, for example, in U.S. Patents 5,399,346 and 5,252,479; and in WIPO publications WO 92/07573, WO 90/06997, WO 89/05345, WO 92/05266 and WO 92/14829, each of which is hereby incorporated herein by reference, in their entirety. These documents provide a description of methods for efficiently introducing nucleic acids into human cells using such retroviral vectors. Other retroviral vectors include, for example, mouse mammary tumor virus vectors (e.g., Shackleford et al., (1988) *PNAS, USA*, 85:9655-9659), human immunodeficiency virus (e.g., Naldini et al. (1996) *Science* 272:165-320), and the like.

Various procedures are also well-known in the art for providing helper cells which produce retroviral vector particles which are essentially free of replicating virus. See, for example, U.S. Patent 4,650,764; Miller, *Human Gene Therapy*, 1:5-14 (1990); Markowitz, et al., *Journal of Virology*, 61(4):1120-1124 (1988); Watanabe, et al., *Molecular and Cellular Biology*, 3(12):2241-2249 (1983); Danos, et al., *PNAS*, 85:6460-6464 (1988); and Bosselman, et al., *Molecular and Cellular Biology*, 7(5):1797-1806 (1987), which disclose procedures for producing viral vectors and

helper cells which minimize the chances for producing a viral vector which includes a replicating virus.

Recombinant retroviruses suitable for carrying out the invention methods are produced employing well-known methods for producing retroviral virions. See, for example, U.S. Patent 4,650,764; Miller, *Human Gene Therapy*, 1:5-14 (1990); Markowitz, et al., *Journal of Virology*, 61(4):1120-1124 (1988); Watanabe, et al., *Molecular and Cellular Biology*, 3(12):2241-2249 (1983); Danos, et al., *PNAS*, 85:6460-6464 (1988); and Bosselman, et al., *Molecular and Cellular Biology*, 7(5):1797-1806 (1987).

Thus, in one embodiment, a recombinant retroviral vector can be utilized to express receptor peptide. Preferably, the retroviral vector will further comprise a regulatory element and exogenous nucleic acid under the control of the regulatory element. Optionally, the retroviral vector can express an antibiotic resistance gene (see Figure 1C). A "covector" can also be utilized to provide a nucleic acid construct comprising the promoter, the regulatory element and exogenous nucleic acid and a second antibiotic resistance gene. The co-vector carrying exogenous nucleic acid also has LTRs modified to promote high-level expression of exogenous nucleic acid (s) only in the presence of the receptor peptide encoded by the recombinant retrovirus and exogenous ligand therefor. Co-infected primary mammalian cells can then be selected using both antibiotics, resulting in a cell population that is dependent on ligand for high-level expression of the exogenous nucleic acid.

The use of invention nucleic acid constructs allows the stable integration and expression of an exogenous nucleic acid(s) or transgene into a wide variety of

cultured cell types. In accordance with another embodiment of the present invention, there are provided recombinant cells containing invention nucleic acid constructs encoding receptor peptides as described herein, optionally further containing regulatory elements operatively associated with exogenous nucleic acid(s). Alternatively, recombinant cells constructs comprising a regulatory element, such as the bombyx mori receptor response element, operatively associated with exogenous nucleic acids, optionally further containing invention receptor peptides.

The amount of exogenous nucleic acid introduced into a host can be varied by those of skill in the art. For example, when a viral vector is employed to achieve gene transfer, the amount of nucleic acid introduced can be increased by increasing the amount of plaque forming units (PFU) of the viral vector.

Suitable means for introducing (transducing) expression vectors containing invention nucleic acid constructs into host cells to produce transduced recombinant cells (i.e., cells containing recombinant heterologous nucleic acid) are well-known in the art (see, for review, Friedmann, (1989) *Science*, 244:1275-1281; Mulligan, (1993) *Science*, 260:926-932, each of which are incorporated herein by reference in their entirety). Exemplary methods of transduction include, e.g., infection employing viral vectors (see, e.g., U.S. Patent 4,405,712 and 4,650,764), calcium phosphate transfection (U.S. Patents 4,399,216 and 4,634,665), dextran sulfate transfection, electroporation, lipofection (see, e.g., U.S. Patents 4,394,448 and 4,619,794), cytofection, particle bead bombardment, and the like. The transduced nucleic acid can optionally include sequences which allow for its extrachromosomal (i.e., episomal) maintenance, or the

transduced nucleic acid can be donor nucleic acid that integrates into the genome of the host.

Exemplary eukaryotic cells suitable for introducing invention expression vectors include, e.g., CV-1 cells, P19 cells and NT2/D1 cells (which are derived from human embryo carcinomas), ES cells (embryonic stem cells), COS cells, mouse L cells, Chinese hamster ovary (CHO) cells, primary human fibroblast cells, human embryonic kidney cells, African green monkey cells, HEK 293 (ATCC accession #CRL 1573; U.S. Patent No. 5,024,939), Ltk$^-$ cells (ATCC accession #CCL1.3), COS-7 cells (ATCC under accession #CRL 1651), DG44 cells (dhfr$^-$ CHO cells; see, e.g., Urlaub et al. (1986) Cell. Molec. Genet. 12: 555), cultured primary tissues, cultured tumor cells, neuronal progenitor or precursor cells, such as *hcn/v-myc* (a tetracycline sensitive retroviral vector selected by G418 resistance), neuronal cells lines such as cerebellum derived neuronal precursors and PC12 cells, neurons, primary astrocytes, oligodendrocytes, and the like. Presently preferred cells include CV-1 and 293 cells.

Invention nucleic acid constructs may be stably incorporated into cells or may be transiently expressed using methods known in the art. Cells are cultivated under growth conditions (as opposed to protein expression conditions) until a desired density is achieved. Stably transfected mammalian cells may be prepared by transfecting cells with an expression vector having a selectable marker gene (such as, for example, the gene for thymidine kinase, dihydrofolate reductase, neomycin resistance, and the like), and growing the transfected cells under conditions selective for cells expressing the marker gene. To prepare transient transfectants, mammalian cells are transfected with a reporter gene (such as the *E. coli* ß-galactosidase gene) to monitor transfection efficiency. Selectable

marker genes are typically not included in the transient transfections because the transfectants are typically not grown under selective conditions, and are usually analyzed within a few days after transfection.

In accordance with a still further embodiment of the present invention, there are provided transgenic animals and methods for producing transgenic animals capable of prolonged and regulated expression of exogenous nucleic acid(s), said method comprising introducing into early-stage embryos or stem cells:

(i) a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element;

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

As used herein, the phrase "transgenic animal" refers to an animal that contains one or more inheritable expression constructs containing one or more exogenous nucleic acid(s) under the transcription control of an operator and/or hormone response element as described herein.

Methods of making transgenic animals using a particular nucleic acid construct are well-known in the

art. When preparing invention transgenic animals, it is preferred that two transgenic lines are generated. The first line will express, for example, a receptor peptide as described above (e.g., VbR). Tissue specificity is conferred by the selection of a tissue-specific promoter (e.g., T-cell specific) that will direct expression of the receptor peptide to appropriate tissue. A second line contains a nucleic acid construct comprising a promoter and exogenous nucleic acid under the control of a regulatory element. Preferably, transgenic animals are produced by the transfection/infection of embryonic stem cells which are employed to produce invention transgenic animals employing methods known to those of skill in art.

In a presently preferred embodiment, an invention transgenic animal contains one or more expression constructs containing nucleic acid encoding receptor peptide and exogenous nucleic acid under the transcription control of a regulatory element. Thus, with tissue specific expression of the receptor peptide as described above and timely hormone treatment, inducible gene expression can be achieved with spatial, dosage, and temporal specificity.

In yet another embodiment of the present invention, there are provided methods of inducing the transcription of an exogenous nucleic acid(s) in a host containing:

(i) a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element;

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor

superfamily which is not normally present in the cells of said host, wherein expression of said receptor peptide is under the control of an inducible promoter, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain; and

(iii) said ligand for said ligand binding domain, wherein said ligand is not normally present in the cells of said host;

said method comprising subjecting said host to conditions suitable to induce expression of said receptor peptide.

In accordance with yet another embodiment of the present invention, there are provided methods for the expression of recombinant products detrimental to a host organism, said method comprising:

transforming suitable host cells with:

(i) a nucleic acid construct comprising a promoter and exogenous nucleic acid(s) which express said recombinant product under the control of a regulatory element; wherein said regulatory element is not normally present in the cells of said host, and

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide

activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain, and

growing said host cells to the desired level in the substantial absence of ligand for said receptor peptide; and

inducing expression of said recombinant product by introducing into said host cells a ligand, which, in combination with said receptor peptide, binds to said regulatory element and activates transcription therefrom.

Recombinant products detrimental to a host organism contemplated for expression in accordance with the present invention include any gene product that functions to confer a toxic effect on the organism. For example, inducible expression of a toxin such as the diptheroid toxin would allow for specific ablation of tissue (Ross et al. *Genes and Development* 7:1318-1324 (1993)). Moreover, the numerous gene products that are known to induce apoptosis in cells expressing such products are contemplated for use herein (see, e.g, *Apoptosis, The Molecular Basis of Cell Death*, Current Communications In Cell & Molecular Biology, Cold Spring Harbor Laboratory Press, 1991). For example, high level expression of 79-amino acid polyglutamine tracts in cells results in apoptosis and rapid cell death.

In accordance with still another embodiment of the present invention, there are provided nucleic acid constructs and cell lines which express a polyglutamine expression cassette (PQEC) under the control of an exogenous ligand to maximize obtaining stable transfected/infected cultured cells which do not undergo

undesired apoptosis. Expression of polyglutamine tracts separately from the huntingtin protein facilitates identification of the toxic properties associated with the polyglutamine tracts themselves. In another embodiment of the present invention, there are provided methods for employing these cell lines to identify novel proteins that could target developing intracellular aggregates (IAs) and either block subsequent growth or facilitate or mark the IAs for degradation.

In accordance with still another embodiment of the present invention, there are provided methods for modulating the transcription of nucleic acid(s) in an *in vitro* system, said method comprising administering to said system an amount of a ligand effective to modulate the transcription of said nucleic acid(s); wherein said ligand is not normally present in said cellular system; wherein said system comprises:

(i) a nucleic acid construct comprising a promoter and said nucleic acid(s) under the control of a regulatory element; and

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain,

In accordance with yet another embodiment of the present invention, there are provided methods for the

treatment of a host in need of gene therapy, said method comprising:

introducing into cells of said host:

(i) a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element;

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein expression of said receptor peptide is under the control of an inducible promoter, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

administering, to said host, ligand for said ligand binding domain.

As used herein, the term "*in vivo* delivery" refers to delivery of biological materials by such routes of administration as oral, intravenous, subcutaneous, intraperitoneal, intrathecal, intramuscular, intracranial, inhalational, topical, transdermal, suppository (rectal), pessary (vaginal), and the like.

As employed herein, the term "ligand" (or ligand precursor) refers to a non-steroidal substance or compound which, in its native state (or after conversion to its "active" form), binds to the receptor peptide, thereby creating a ligand/receptor peptide complex, which in turn can bind an appropriate response element and activate

transcription therefrom. Ligands function to modulate transcription of nucleic acid(s) maintained under the control of a response element. In accordance with one aspect of the present invention, unless and until a suitable ligand is administered to the host, substantially no transcription of the desired exogenous nucleic acids occurs.

Preferred ligands contemplated for use in the practice of the present invention are ligands characterized as not normally present in the cells of the host to be treated. Such ligands are referred to as being exogenous to the host. An example of a class of ligands not naturally present in mammalian systems are compounds referred to as hydrazines, preferably diacyl hydrazines.

Hydrazines contemplated for use in the present invention include compounds which are readily available and are/or relatively inexpensive to manufacture. One such compound, tebufenozide, is a non-steroidal ecdysone agonist which is used commercially as an insecticide. This compound specifically targets lepidopteran species, including *bombyx mori*. Tebufenozide has undergone extensive testing in animal hosts and has proved to be of very low toxicity to mammals and other non-insect species.

Exemplary hydrazines contemplated for use herein are mimics of the naturally occurring ecdysones, i.e., synthetic organic compounds which have transactivation activity characteristic of the naturally occurring ecdysones. Examples of such compounds, referred to herein as ecdysone mimics, include 1,2-diacyl hydrazines (e.g., tebufenozide and others described in U.S. Patent Nos. 5,424,333 and 5,354,762, the entire contents of each of which are hereby incorporated by reference herein), N'-substituted-N,N'-di-substituted hydrazines (e.g., those

described in U.S. Patent No. 5,117,057, the entire contents of which are hereby incorporated by reference herein), dibenzoylalkyl cyanohydrazines (e.g., those described in European Application No. 461,809, the entire contents of which are hereby incorporated by reference herein), N-substituted-N-alkyl-N,N'-diaroyl hydrazines (e.g., those described in U.S. Patent No. 5,225,443, the entire contents of which are hereby incorporated by reference herein), N-substituted-N-acyl-N-alkyl, carbonyl hydrazines (e.g., those described in European Application No. 234,944, the entire contents of which are hereby incorporated by reference herein), N-aroyl-N'-alkyl-N'-aroyl hydrazines (e.g., those described in U.S. Patent No. 4,985,461, the entire contents of which are hereby incorporated by reference herein), and the like.

Since it has been previously reported that the above-described diacyl hydrazines are neither toxic, teratogenic, nor known to affect mammalian physiology, they are ideal candidates for use as inducers in cultured cells and transgenic mammals according to invention methods.

Ligands are administered in a manner compatible with the route of administration, the dosage formulation, and in a therapeutically effective amount. The required dosage will vary with the particular treatment desired, the degree and duration of therapeutic effect desired, the judgment of the practitioner, as well as properties peculiar to each individual. Moreover, suitable dosage ranges for systemic application depend on the route of administration. It is anticipated that dosages between about 10 micrograms and about 1 milligram per kilogram of body weight per day will be used for therapeutic treatment.

An effective amount of ligand contemplated for use in the practice of the present invention is the amount of

ligand (e.g., diacyl hydrazine(s)) required to achieve the desired level of transcription and/or translation of exogenous nucleic acid. A therapeutically effective amount is typically an amount of a ligand or ligand precursor that, when administered in a physiologically acceptable composition, is sufficient to achieve a plasma concentration of the transcribed or expressed nucleic acid product from about 0.1 μg/ml to about 100 μg/ml, preferably from about 1.0 μg/ml to about 50 μg/ml, more preferably at least about 2 μg/ml and usually 5 to 10 μg/ml.

Ligand can be administered in a variety of ways, as are well-known in the art, i.e., by any means that produces contact between ligand and receptor peptide. For example, such ligands can be administered topically, orally, intravenously, intraperitoneally, intravascularly, and the like. The administration can be by any conventional means available for use in conjunction with pharmaceuticals, e.g., by intravenous injection; either as individual therapeutically active ingredients or in a combination with other therapeutically active ingredients. Ligands contemplated for use in the practice of the present invention can be administered alone, but are generally administered with a pharmaceutical carrier selected on the basis of the chosen route of administration and standard pharmaceutical practice.

Therapeutic compositions containing suitable ligand are preferably administered intravenously, as by injection of a unit dose, for example. The term "unit dose," when used in reference to a therapeutic composition of the present invention, refers to a quantity of ligand suitable as unitary dosage for the subject, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent, i.e., carrier, or vehicle. It may be

particularly advantageous to administer such compounds in depot or long-lasting form as discussed hereinafter.

Suitable regimes for initial administration and booster shots are variable, but are typified by an initial administration followed by repeated doses at one or more intervals by a subsequent injection or other administration. Alternatively, continuous intravenous infusion sufficient to maintain concentrations in the blood in the ranges specified for *in vivo* therapies are contemplated.

In accordance with still another embodiment of the present invention, there are provided methods for the treatment of a host in need of gene therapy, said method comprising:

introducing into cells obtained from said host:

(i) a nucleic acid construct comprising a promoter and exogenous nucleic acid(s) under the control of a regulatory element; and

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain; and

to provide modified cells;

reintroducing the modified cells into said host; and

administering, to said host, ligand for said ligand binding domain.

In accordance with this embodiment of the present invention, the exogenous nucleic acid is introduced directly into cells obtained from a donor (host or separate donor). Such cells are then implanted within the host. In a presently preferred embodiment, the transplanted cells are autologous with respect to the host. Autologous means that the donor and recipient of the cells are one and the same.

The concept of gene replacement therapy for humans involves the introduction of functionally active "wild type" or "therapeutic" nucleic acids into the somatic cells of an affected host to correct a gene defect or deficiency. However, in order for gene replacement therapy to be effective, it must be possible to control the time and location at which gene expression occurs.

Genes that encode useful "gene therapy" proteins that are not normally transported outside the cell can be used in the invention if such genes are "functionally appended" to a signal sequence that can "transport" the encoded product across the cell membrane. A variety of such signal sequences are known and can be used by those skilled in the art without undue experimentation.

The invention will now be described in greater detail by reference to the following non-limiting examples.

Example 1
Construction of Lepidopteran-derived fusion receptor peptide

The VP16 t-domain (activation domain) is fused in frame onto *bombyx mori*-derived nuclear receptor (bR) at an internal MluI site near the N-terminus (VbR, Figure 1A). An AscI site immediately downstream of the ATG start codon allows multimerization of VP16 domains for VbRs with multiple VP16 activation domains. Note that unlike Drosophila ecdysone receptor (dEcR), bR has virtually no C-terminal domain downstream of the LBD. Several amino acids of the hormone binding domain are replaced with sequences from the retinoic acid receptor or the thyroid hormone receptor (VbRRA and VbRTR, respectively) (Figure 1B). These mutant variations have properties slightly different from VbR, and in addition have silent mutations that eliminate BclI and BstBl sites in VbR. These are useful in the construction of many VbR-based retroviral constructs.

Example 2
Construction of multifunctional regulatory proteins

The small size and simplicity of the VbR system lends itself to the development of multifunctional regulatory proteins which are functional on both hormone receptor and TTA responsive promoters. VbR fusion to the tetracycline transactivator results in a hybrid protein called TTMT (for Tebufenozide/Tetracycline Modulated Transactivator; see Figure 1C). This protein functions jointly as a ligand-modulated regulator of gene expression from both tetO and EcRE (ecdysone response element)-containing promoters either separately or simultaneously. This compact protein, encoded in approximately 3 kb, confers

constitutive activation of promoters containing tetracycline operators (tetOs) in the absence of tetracycline/analogs. In the presence of tet analogs, the tetO-binding function of TTMT is blocked, deactivating responsive promoters. The VbR half of the chimeric protein is constitutively inactive on EcRE-containing promoters, but may be stimulated to transactivate responsive promoters to a high level by the addition of muristerone A (murA) or tebufenozide (teb). Through such a dual system, two separate responsive promoters may be simultaneously regulated by two separate ligands through a single protein. Additional variants using the reverse-Tet-transactivator, other VbR variants, novel response elements, and other regulatory proteins, will presumably result in further customized variations on this same theme.

A single transgenic animal generated with the TTMT protein would be simultaneously responsive to transgene regulation of two separate transgenes by two separate ligands. Further, individual promoters can be produced that respond to both halves of the protein, allowing the transgene to be regulated by both ligands, at the discretion of the investigator.

Example 3
Transient transfection assay

Transient transfection experiments are carried out in CV-1 cells or 293 cells using modified drosophila ecdysone receptor (VdEcR) or modified bombyx hormone receptor (VbR) with or without dimer partners. Luc activity is examined 40 hours after transfection and stimulation with ligand (Figure 2A). Both receptors are assayed in CV-1 cells with 1 μM MurA and no exogenous dimer partner. Note that bR responds, while dEcR does not. Both receptors are also assayed in the presence of 5 μg/ml tebufenozide (Figure

2B). Note that VbR responds with greater than 200-fold induction while drosophila EcR does not respond at all, and in fact, decreases slightly. Both receptors are also assayed with and without dimer partners. VbR is stimulated with 5 μg/ml tebufenozide and VdEcR is stimulated with 1 μM MurA. Note that RXR and Usp are necessary for VdEcR while VbR is constitutively activated by the presence of exogenous dimer partner and is inhibited by ligand addition. VbR, with no exogenous dimer partner, has an extremely low base-line which is stimulated 200-300 fold by ligand. Note also from this experiment that VdEcR tends to have a much higher baseline than VbR without dimer partner.

A side-by side comparison of 2.5 μg/ml teb and 1 μM MurA on VbR show both ligands are effective at stimulating VbR. Tebufenozide is administered to cells co-transfected with nucleic acids encoding receptors: LINX (which encodes the tet-transactivator (TTA)), VbR, TTMT (which encodes the TTA-VbR fusion protein, and TTMT-2V which is like TTMT, but with two VP16 t-domains) to assay for EcRE-Luc activity (Figure 2C). Note that teb stimulates VbR, and the TTMT fusions, but has no action on TTA (LINX). The same experiment as that described for 2C, only with a TetO-Luc reporter, which responds only to TTA or fusion proteins is summarized in Figure 2D. Doxycycline (dox) acts to block constitutive activity of TTA. Note that in this experiment, TTMT and TTMT-2V work better than TTA only (LINX) to stimulate the TetO-luc promoter/reporter. They are also efficiently blocked by dox. An experiment on CV-1 cells transfected with TTMT and two reporters (6TetO-luc and 4-EcRE-βgalactosidase) shows both promoters could be simultaneously influenced by the presence or absence of both ligands (Figure 2E).

Example 4

Ligand mediated transgene regulation

Transgene regulation by teb or MurA in stably infected rat fibroblasts is assayed utilizing recombinant retroviral constructs as the reporter, and VbR encoding retroviral vectors to provide the receptor. Bulk infected selected populations stimulated with 2.5 μg/ml teb or 1 μM MurA indicate b-gal or tyrosine hydroxylase infected cells. Quantitative analysis of an individual VbR-luc clonal fibroblast line, and a VbR-GH clonal line show that 0.25 μg/ml teb is as effective at stimulating full activity as higher levels (Figure 3).

Example 5

Single-plasmid retroviral constructs

The plasmid, pBO, utilizes a responsive internal promoter to regulate transgene expression and also contains a VbR-responsive 3' LTR to autoregulate the VbR receptor itself (Figure 4). The plasmid, pNA, is constructing using a 3'-responsive LTR to regulate transgene expression. pBO utilizes an EcR responsive internal promoter and a TTA responsive 3' LTR. Both are host to regulation by the TTMT protein, simultaneously. The purpose of this design is to provide chronic stimulation of the duplicated 5' LTR along with regulated expression of the internal EcR responsive promoter. By removing all of the native promoting sequences from the 6To LTR with the exception of the TetO's, variants of this construct are also inactivatable.

Example 6

Construction of plasmid vectors

VbR and TTMT systems including constitutive and inducible promoters are integrated into a modular plasmid

vector known as "pBW" (Figure 5). The purpose of pBW is to integrate all of the elements of bombyx receptor- or tetracycline-analog responsiveness into a single plasmid vector destined for use in producing transgenic animals by methods such as pronuclear injection. pBW is to facilitate construction of transgenic animals or stably transfected cells using VbR variants. The pBW design simplifies the insertion of responsive promoters, transgenes, constitutive promoters, and the various VbRs, TTMTs, and related regulatory proteins. With pBW, the transgene regulatory properties of generated animals should be more predictable and reproducible from founder to founder than multi-vector systems.

There are two major components of this system: the first is a tissue-specific expression cassette to produce the receptor and/or antibiotic resistance genes. The second is an expression cassette with a responsive promoter and a transgene. Polyadenylation signals flank all of the expression cassettes to provide efficient polyadenylation (p(A)) of transcribed RNAs. The use of rare restriction endonuclease sites within strategic locations and polylinkers of pBW will ease cloning of transgenes. An adjunct shuttle vector derived from pBSK (Stratagene) called SKSP contains rare sites flanking the polylinker and compatible with pBW. pBW also shares a number of sites with pBO to further simplify construction. pBW is constructed within the plasmid pcDNA3 (Invitrogen).

Example 7
Regulating transgene expression in target cells *in vitro*

Individual desired elements can be inserted sequentially into a recombinant retrovirus. Recombinant retrovirus starts as a polylinker composed of 5' *Not*I-*Mlu*I-

*Nru*I-*Eco*RI-*Asc*I-*Pml*I-*Bst*BI-*Bam*HI-*Hin*dIII-*Hpa*I-*Cla*I-*Nsi*I-*Kpn*I 3' sites. The retroviral Ψ gln (see Adam and Miller in *J. Virol.* 62:3802 (1988) and Barklis et al., in *Cell* 47:391 (1986)) is inserted into sites of the recombinant retroviral polylinker. The internal CMV promoter is inserted *Bam*HI-*Cla*I into the *Bam*HI-*Cla*I sites of the evolving recombinant retroviral vector. All LTRs destined for insertion into the 5' location are produced by low-cycle/high-fidelity PCR production and end primers with compatible *Not*I-*Mlu*I sites for insertion into the *Not*I-*Mlu*I sites. 3' LTRs are generated with *Cla*I-*Nsi*I compatible ends for insertion into these sites. The R region of the 3' LTR is included in the downstream primer sequence. Transgenes, receptor cDNAs, or selectable marker genes are inserted into the remaining polylinker from the *Eco*RI site through the *Bam*HI site, or the remaining polylinker after the internal promoters (*Hin*dIII-*Cla*I).

The resulting vectors are characterized by restriction digests and mapping, amplified by large-scale plasmid preparation, and prepared for transfection into packaging cells by standard methods. Transient retroviral production in 293 and 293T cells has been previously described (see Pear et al., in *PNAS* 90:8392 (1993)). Forty-eight hours after transient retroviral production, the conditioned media is removed, filtered through 0.45 mm filters, and frozen at -70°C until use. 10-cm dishes of primary-cultured Fisher rat abdominal fibroblasts (Rosenberg et al., *Science* 242:1575 (1988)) at approximately 50% density are infected with 1/10 volume of virus-containing media and 8 μg polybrene/ml for 48 hours. Infected cells are then selected under the appropriate antibiotic until non-infected cells are cleared from the population (approximately 10 days). Resistant colonies are trypsinized, pooled and passaged. The resulting population

is then infected with the second virus, and reselected again in media containing both antibiotics. The infection and selection process for the second round of viral infection proceeds as described for the first round. The final population of doubly resistant primary fibroblasts are then pooled and passaged to 6 or 24-well Costar plates for assay *in vitro*.

The short (24 hour) course is performed in 24-well Costar dishes with cells at approximately 80% confluency. 5 μl of ligand (final concentration of 1 μM) or vehicle (20% EtOH) is added to the 1 ml of media in each well. Eight hours after first ligand administration, the cells are all washed in PBS and the media is replaced. β-galactosidase histochemical reaction is performed on 1.5% glutaraldehyde-fixed cells essentially as described by Shimohama et al. in *Brain Res Mol Brain Res* 5:271 (1989) in a 37°C environment for 2 hours. The 4-day time course is performed in an identical fashion, except that the cell population is plated at only 20% confluency at the start of the experiments (and grow to near-total confluency by the end of the 4 day period) and neither ligand nor media is changed or replaced.

For long-term culture experiments, MMBG fibroblasts are plated in triplicate at high density in 24-well plates. Within 48-72 hours of plating, when the cells reach 100% confluency within the wells, 1 μM ligand or vehicle is added to one group of plates, while others are left completely untreated. The wells are then placed in a 37°C-10% CO_2 environment for the next 25 days without any media changes. On day 23, the plates that had received neither vehicle nor ligand on day 1 are removed and half of the wells are given vehicle and the other half are given 1 μM ligand for 40 hours.

After 25 days of nutrient deprivation and contact inhibition, all of the plates are processed for βgal histochemical staining as above.

MMGH fibroblasts are produced by infecting a fibroblast line with the recombinant retroviral vectors containing desired elements. Cells are selected under both antibiotics as described previously. After preliminary examination of ligand-induced hGH production in the bulk population, the MMGH population is plated at high density in 6-well Costar plates. Forty-eight hours after the initial plating and when the cells are essentially completely growth inhibited by contact, the medium of all plates is replaced with DME containing 2% FBS. Under these conditions, primary rat fibroblasts stop dividing even if they are not confluent, and settle out into a distended morphology with prominent nuclei characteristic of severely growth-arrested fibroblasts. Six days after acclimatization to this culture environment, 1 ml of medium is harvested from each well and plate of the experiment and stored frozen at -20°C. The remaining unused media is discarded at the same time of harvest and replaced with fresh DME-2% FBS. Groups of wells are treated or not with 1 μM ligand. Hormonally treated wells are washed with medium to remove as much residual ligand as possible prior to medium replacement when ligand treatment is discontinued. This routine is repeated daily for the 21-day extent of the experiment.

After all time-point samples have been collected and frozen at -20°C, they are simultaneously processed for the presence of the hGH protein by ELISA (Boehringer Mannheim) following the protocol recommended by the manufacturer. ELISA data are quantitated on an MR700 microplate reader (Dynatech Laboratories, Chantilly, VA) and compared to a standard curve generated using purified hGH to determine

picogram GH amounts. Control wells are TH (tyrosine hydroxylase) producing fibroblasts, but display all of the characteristics of the recombinant retrovirus except for the hGH transgene and secreted protein.

Example 8

Expression Vectors and Transient Transfection Assays

All receptors were subcloned into vector LNCX (A.D. Miller, GenBank accession no. M28247, with an extended polylinker) for use in transient transfection assays. VBR (or CVBR) when referring to the retrovirus) was produced by insertion of VP16 (14) sequences with a synthetic ATG start codon into the amino-terminal region of BE up to, and in-frame with, the *Mlu*1 site (corresponding to amino acid 26). VP16 primers: (5') 5'-GAGAGAAGCTTATGGCGCGCCCGACCGATG (SEQ ID NO:5) and (3') 5'-CACACACGCGTGTACTCGTCAATTCCAAG (SEQ ID NO:6). VDE (or CVDE) was produced in a similar fashion by fusing VP16 sequences in-frame at a novel *Nco*I site corresponding to amino acid 68. The downstream primer was made *NcoI* compatible. Plasmid template (100 ng), each primer (500 ng), and reaction conditions outlined by the manufacturer for Pwo (Boehringer Mannheim) high fidelity polymerase, were used with a program of 1 min at 94°C, 1 min at 45°C, and 1 min at 72°C for 20 cycles for production of all PCR products used herein. LNCX-Usp and RXR were produced from *Eco*RI fragments encoding this complete cDNAs, filling with the Klenow fragment of DNA polymerase, and inserting the fragments into the *Hpa*I site of LNCX. Orientation was determined by restriction endonuclease digestion. Transient transfection in CV-1 and 293 cell was performed by using standard methods (Sambrook et al (1989) Molecular Cloning: A Laboratory Manual 2nd edition) in triplicate in 24-well plates by calcium phosphate coprecipitation of 100-ng receptor(s), reporter

plasmid E4-luc, and pCH110 as an internal control. Briefly, E4-luc is four tandem EcREs inserted upstream of a thymidine kinase gene minimal promoter directing luciferase expression EcRE oligonucleotides were as described (Thomas et al. (1993) Nature 362:471-475) with *BamHI-Bg/II* compatible ends. Ligand (1 μM) was added at the time of transfection, and 40 h later cells were lysed and extracts were used for β-galactosidase assay and measurement of luciferase activity in an analystical bioluminescence photometer.

CVDEiR and MS vectors used only as retroviruses in production of stable cell lines were produced as follows: The VDEiR expression cassette was constructed by modifying the ATG start codon of human RXRα to contain an overlapping *BstXI* site for fusion into the ATG start codon of the 0.8-kb EMCV IRES (Jang et al (1989) J. Virol 63:1651-1660 and Hoshimaru et al (1996) PNAS 93:1518-1523) sequence. The IRES-RXR cassette was then subcloned downstream of VDE in LNCX-ofVDE (CofVDE), a VDE variant vector with lower basal transactivation levels in superphysiological RXR environments, to produce CVDEiR. The MS retroviral reporter was produced by removing the Moloney murine leukemia virus enhancer core from the 3' long terminal repeat (LTR) with *NheI-XbaI* digestion and replacing this sequence with four tandem EcREs, resulting in a reporter analogous to the E4-luc reporter plasmid.

Example 9
Gel Mobility Shift Assays

Gel mobility shift analysis was performed by *in vitro* translation of receptor ORFs subcloned into pBSK (Stratagene), pGEM-3 (Promega), and PSL301 (Invitrogen) in the presence of [^{35}S]Met by using T3/T7 TNT (Promega)

transcription/translation and following the manufacturer's protocol. *In vitro* translated proteins were qualitatively examined by 5% SDS/PAGE with protocols as described elsewhere (Sambrook et al (1989) Molecular Cloning: A Laboratory Manual 2nd edition) and quantified by PhosphorImager (Molecular Dynamics) exposures of dried gels. Amounts of proteins used in gel shift assays were normalized by using quantitative data and correction for predicted Met residues in individual constructs. Double-stranded EcRE probes corresponding to response elements described above were labeled by filling of the Klenow fragment of DNA polymerase with [^{32}P]dCTP and unlabeled dGAT by standard methods. Reaction conditions for protein-probe interaction and gel electrophoresis were essentially as described by Yao *et al.* ((1992) Cell 71:63-72) except, to facilitate comparison between samples, reaction mixtures (including dimer partners) and probe) were prepared as a mixture and distributed equally to individual tubes with receptor proteins. The reactions were allowed to proceed at 23°C for 5 min. at which time ligand or vehicle was added and the reaction allowed to continue for an additional 20 min.

Example 10

Construction of *bombyx mori* receptor (BE) - *drosophila melanogaster* ecdysone receptor (DE) Chimeric Receptors

Chimeric receptors were produced by taking advantage of conserved DNA sequences and sites within the receptors and by introducing compatible sites by low cycle, high fidelity PCR (described above) of *bombyx mori* receptor (BE) and *drosophila melanogaster* ecdysone receptor (DE) templates to produce fusion proteins. DEBE-A/B was produced by PCR and replaced the BE /AB domain (amino acids 1-197, as set forth in SEQ ID NO:3) with amino acids 1-255

of DE introducing a novel *Apa*1 restriction site. DEBE-C replaced the A/B and C domains (amino acids 1-273, as set forth in SEQ ID NO:3) of BE with amino acids 1-331 of the DE protein by introduction of a compatible *Eag*1 site into BE at the fusion junction. DEBE-D replaced BE from amino acids 1-363, as set forth in SEQ ID NO:3, with amino acids 1-430 of DE by *Kpn*I digest. BEDE was the reverse of this construct and replaced amino acids 1-430 of DE with amino acids 1-363 (as set forth in SEQ ID NO:3) of BE by using *Kpn*I partial digest. DEBH was produced from DEBE-C by removing all BE sequences except the D domain by *Kpn*I digestion and inserting sequences encoding the intact E and F domains (amino acids 430-878) of the original DE after partial *Kpn*I digestion. VEH was produced by excision of the central region of DEBH, including the heterologous hinge domain, and inserted into VDE, replacing the analogous region.

Example 11
Construction of E Domain Chimeric Receptors

The chimera, BEDB, was produced by insertion of the 654-bp *Kpn*I-*Bg*/II fragment encoding precisely the DE hormone-binding domain into the identical analogous sites in BE. Two unique sites within the hormone-binding domain of BEDB were used in the production of the chimeras. A unique *Aat*II site lies approximately 120 bp downstream of the 5' *Kpn*I site, and a unique *Eag*I sites lies approximately 440 bp downstream. PCR was used as described above, and corresponding fragments of BE with appropriate compatible ends were generated and subcloned into BEDB digested with *Kpn*I + *Aat*II(BKE), *Aat*II + *Eag*I (BAE), *Aat*II + *Bg*/II (BAB), and *Eag*I-*Bg*/II for BEB. The resulting constructs were translated, examined by SDS/PAGE,

quantified, and normalized as described above. Conditions for gel mobility shift assay were also as described above.

Example 12
Infection of Cells with Recombinant Retroviruses

Transient restroviral production in 293 cell types has been described (Pear et al. ((1993) PNAS 90:8392-8396). Forty-eight hours after transient retroviral production, the conditioned medium was removed, filtered through 0.45-μm filters, and frozen at - 70°C until use. Ten-centimeter dishes of primary-cultured Fischer rat abdominal fibroblasts (Schinstine et al. (1992) Neurochem 58:2019-2029) and 293 and CV-1 cells at approximately 50% density were infected at a multiplicity of infection of approximately 0.05 with virus-containing media and 8 mg/ml Polybrene for 48 h. Infected cells were then selected with 4-6 mM L-histidinol until noninfected cells were cleared from the population: approximately 10 d for FF12s, 21 dafor 293s, and >30 d for CV-1s (CV-1 cell populations were resistant to rigorous selection with L.-histidinol). Resistant colonies from multiple plates were trypsinized, pooled, and passaged. The resulting bulk population was then split and infected with receptor-bearing virus and reselected again in medium containing both L-histidinol and 400-800 μg/ml G418. The infection and selection process for the second round of viral infection proceeded as described for the first round. The final population of doubly resistant primary fibroblasts was then pooled and passaged in triplicate in 6- or 24-well Costar plates for β-galactosidase staining by standard methods or luciferase assay.

Example 13

CV-1 Cells Support BE-Mediated Transactivation Without Exogenous Dimer Partner

Both receptors were tested to determine the effect that differential endogenous dimer partner availability had on transactivation of responsive promoters. Relative luciferase activity employing the reporter plasmid, E4-luc, were performed from VDE and VBR transiently cotransfected into 293 or CV-1 cells with either Usp, RXR, or no added heterodimer partner. Luciferase activity was assayed in the presence of either vehicle, 1 μM murA, or 1 μM tebufenozide.

In the presence of an equimolar exogenously added dimer partner, VDE and VBR transactivation was similar in both cell types. With Usp, both proteins were induced less than 2-fold by ligand and displayed a high level of basal transactivation. With RXR, VDE displayed an average relative 5.35-fold induction across both cell types, whereas VBR was induced only 2.35-fold, even though the absolute level of induction matched or exceeded the expression with VDE. The decreased relative induction of VBR + RXR resulted from approximately doubled basal activity levels compared with VDE + RXR. With no exogenous dimer partner, both receptors in both cell types exhibited dramatic 15- to 80-fold decreases in basal transactivation.

With no exogenous dimer partner, VDE failed to respond to tebufenozide, whereas VBR continued to respond well to both murA and tebufenozide. The addition of RXR to VDE-transfected 293 cells increased the maximum murA-treated expression level by only 20%, indicating that the high level of endogenous RXR in 293 cells is near saturation for heterodimerization with DE. By comparison,

CV-1 cells supported murA stimulation of VDE at only 13% of the expression level of VDE with added RXR (Fig. 1B). VBR was active at levels similar to VDE in the high RXR background of the 293 cells; however, in the CV-1 cells, tebufenozide-stimulated VBR exhibited both the highest absolute level of transactivation and the greatest relative induction (160.2-fold) of any of the other combinations of receptors and ligands tested. In the CV-1 cells, tebufenozide-stimulated VBR displayed 21-fold greater relative induction and an absolute expression level 9.25 times the level of VDE treated with mutA.

Example 14
BE Displays Facilitated Heterodimer Formation with Usp and RXR

The high level function of VBR in the low heterodimer partner environment of the CV-1 cells suggested that BE might be capable of efficient function with lower levels of dimer partner than DE. To explore this possibility, cell-free experiments with gel mobility shift analysis and equivalent amount of *in vitro* translated BE and DE in combination with Usp or RXR dimer partners and both ligands were performed. Neither protein alone formed a dimeric complex with the EcRE probe unless either the Usp or RXR dimer partner proteins were added. In the presence of Usp, DE formed a comparatively weak interaction with the EcREs that was increased approximately 4-fold by murA and 2-fold by tebufenozide. In contrast, BE produced a strong shift with Usp that was only slightly influenced by either ligand. In the absence of any ligand, BE bound probe 5 times more efficiently than DE. With ligand, the BE probe shift was observed to still exceed the DE shift and was double the DE + Usp band volume. Both receptors combined with RXR displayed an absolute dependence on ligand for heterodimer formation; however, whereas BE + RXR with

either ligand displayed a prominent shifted band, an equivalent amount of DE resulted in a barely detectable shifted band with murA and no detectable band shift with tebufenozide. The BE + RXR tebufenozide- and murA-induced shifts were determined to be >15 times the DE + RXR shift with murA.

Example 15

BE-DE Chimeras Reveal That the D Domain Is Critical to High Affinity Heterodimerization

Gel mobility shift assays indicated that BE had higher affinity heterodimer formation than DE. To further define the subdomains and molecular determinants resulting in the BE high affinity phenotype, PCR mutagenesis and internal shared restriction endonuclease sites were used to produce BE-DE chimeric receptors that were assayed for their ability to bind the EcRE probe with either Usp or RXR and with or without ligand. Because previous results indicated that murA stimulated both EcRs, for purposes of direct comparison only murA was used for gel shift analysis of the chimeras. Fusion constructs began at the N terminus of BE and sequentially replaced the A/B (DEBE-A), C (DEBE-C), and D (DEBE-D) domains with DE sequences. BEDE was the reverse, encoding BE up to the D-E domain boundary with the DE E and F domains. DEBE-A and DEBE-C with Usp shifted nearly as efficiently as native BE. The absence of shift for DEBE-D, however, and the substantial shift seen for BEDE + Usp suggested that the high affinity determinant for BE-Usp heterodimerization resided in the D, or hinge domain. The same constructs in combination with RXR revealed a different shift pattern. BE and DEBE-A both responded strongly to ligand to bind and shift the probe. DEBE-C with ligand was approximately 40% decreased relative to native BE; however, it still displayed 16-fold greater binding than native DE. DEBE-D had lost all high affinity

DNA-binding complex formation and functioned indistinguishably from the native DE with or without ligand. MurA-stimulated BEDE + RXR produced a band shift 4 times the intensity of the native DE + RXR stimulated with murA, but BEDE was clearly significantly impaired to heterodimerization with RXR compared with native BE and other higher affinity chimeras like DEBE-A and DEBE-C.

For further confirmation of the role of the BE D domain in dimer partner affinity, the D region of DE was replaced with the BE D domain to produce DEBH. Native DE and DEBH were compared side by side for binding to EcREs with both Usp and RXR dimer partners. Gel mobility shift of *in vitro* translated and normalized DE and hinge-substituted DEBH was performed, in the presence of either Usp or RXR, and further in the presence of vehicle or 1 μM murA. Compared with DE + Usp, DEBH + Usp averaged 5-fold greater probe binding with murA and 9-fold greater binding without ligand. DEBH + RXR exhibited both decreased probe shift in the absence of ligand and increased probe binding with murA (approximately 3-fold over native DE) for a 10-fold relative induction by ligand compared with 2.5-fold for DE. Transient transfection analysis of a VP16-DEBH construct (VEH) in CV-1 cells revealed that VEH shared characteristics of both DE and BE.

VEH displayed the high basal transactivation level and low relative induction (average of 2.5 fold) characteristic of BE with superphysiological RXR. VEH displayed 5- to 8-fold greater basal and murA-activated expression than native DE in the absence of any exogenous dimer partner. MurA-stimulated VEH expression exceeded even the expression level for murA-stimulated BE by nearly 2-fold. As predicted, without the BE hormone-binding domain, DEBH did not to respond to tebufenozide. Additional chimeric proteins with a subset for BED-region sequences did not

function as well as DEBH with replacement of the entire D domain, suggesting that multiple discrete determinants presumably over much of the hinge region contribute to the high affinity phenotype.

Example 16
Discrete Functional Regions Within the E Domain Revealed by BE-DE Chimeras

Although the preceding data reveal that the BED domain alone was sufficient for high affinity heterodimerization with Usp, a comparison of the relative shift of the DEBE-D and BEDE chimeras with both partners suggested that additional determinants within the BE hormone-binding E domain contributed to a high affinity interaction with RXR. To further study the BE E domain, chimeric proteins were produced by using BEDE as a template. Five E-domain chimeras named BEDB, BKE, BAE, BAB, and BEB were prepared. BEDB is identical with BE with the exception of complete replacement of the E domain with corresponding sequences from DE, BEDB was identical with BEDE with the exception of replacement of the large C-terminal DE F domain for the 20-amino acid F domain of BE. BKE, BAE, BAB, and BEB are derived from BEDB and are sequential, C-terminal moving replacements of three regions within the DE-derived E domain. Gel mobility shift assays with BE, BEDE, and BEDB revealed that BEDE and BEDB are essentially identical in their properties of complex formation, suggesting that the F domain does not play a significant role in heterodimerization and DNA binding. In addition, BEDB recapitulates the high affinity USP binding and the low affinity RXR binding of BEDE.

The BEDB-derived chimeric receptors reveal several functional subdomains over most of the BE E domain. By using unique internal sites, the E domain was subdivided

into approximately thirds called E_1 E_2 and E_3. Treatments were with vehicle (-), 1 μM murA (M), or 1 μM tebufenozide (T). Chimera BKE with replacement of the DE E_1 and E_2, regions and chimera BAE with replacement of only the E_2 region displayed very similar patterns of shift. Both chimeras were significantly impaired in complex formation with Usp and RXR relative to the original BEDB chimera, suggesting that there were fundamental incompatibilities between subdomains of the *Bombyx* and *Drosophila* E domains. The similarity of these two constructs suggested that DE and BE E_1 regions were likely to be functionally similar, but that the E_2 regions had very different properties. Curiously, although both basal and murA-stimulated heterodimerization were inhibited, BKE/BAE-Usp heterodimerization was significantly stimulated by tebufenozide, indicating that the tebufenozide response determinant is within E_2 but that the high affinity RXR-binding determinant lies elsewhere.

When E_2 was combined in tandem with E_3 in the chimeric receptor BAB, high affinity heterodimer formation indistinguishable from native BE heterodimerization was observed with both Usp and RXR. To determine whether the E_3 region alone could confer high affinity heterodimerization with RXR, the chimera BEB, with substitution of only the E_3 region, was produced and tested. BEB displayed a pattern of response essentially identical with the original BEDB with low-level interaction with RXR and only upon stimulation with murA. BEB, which lacked the E_2 region of native BE, was not responsive to tebufenozide. Taken together, these results indicate that the BE D domain and E_3 region, in combination with ligand-binding determinants within the E_2 region, function in concert to produce high affinity, ligand-dependent heterodimerization between BE and the RXR protein.

Example 17

Comparison of VBR and VDE Function in Stably Transduced Cells

Building on transient transfection and gel shift data, DE an BE variants were directly compared for their ability to function at low or single copy in stably transduced cell types. In addition to direct comparison of CVDE and CVBR retroviruses, an RXR-bearing VDE variant called CVDEiR was also tested that previous studies had determined to provide the best combination of low basal expression and relatively high induction for VDE cotransferred with supplemental RXR. These three receptor-bearing vectors and an empty LNCX control were used to infect 293, CV-1, and primary fibroblast cells harboring an EcRE-responsive retroviruses termed MS. MS encoded a 3' LTR with the core enhancer elements replaced with four tandem EcREs and an internal simian virus 40 promoter vector directing resistance to the drug L-histidinol. Parallel MS vectors encoding either the LacZ gene (MS-Z) or a luciferase transgene (MS-luc) were used in this comparative study, the former as an indication of the number of responding cells and the latter for quantitative purposes.

Three cell types, 293s, CV-Is, and rat primary fibroblasts (line FF12), were selected for assay of reporter/receptor virus combinations based on their relative capacities to supports high level DE-mediated transactivation. The human 293 cell line support essentially full DE activation with no requirement for added RXR, whereas CV-1s require the contransfection of exogenous dimer partner for DE function. FF12s were selected as the third cell type because they are refractory to chemical methods of transduction and reliant on the use of biological vectors such as recombinant retroviruses for efficient gene transfer. Furthermore, primary fibroblasts

are frequently used as autologous donor cells for transplantation approaches to somatic gene therapy.

All three cell types were infected with both the reporter and receptor-bearing vectors and treated for 72 h with either vehicle, 1 μM murA or 1 μM tebufenozide. Histochemical staining of MS-Z infected cells indicated that all three receptor types were capable of high level ligand-responsive transactivation in coinfected 293 cells. DE-derived constructs responded well to murA but were unable to respond strongly to tebufenozide. CVBR, on the other hand, responded moderately to murA but strongly to tebufenozide. MS-LacZ cell types infected with an empty LNCX vector displayed no staining when stimulated with any ligand and further confirmed observations from vehicle-treated cells that the MS vector permitted only relatively low basal levels of transgene expression. Approximately 30% of CVDE-, CV-DEiR-, and CVBR-infected MS-Z 293 cells responded to murA. Tebufenozide stimulated <1% of CVDE-infected cells and <4% of CVDEiR-infected MS-Z 293s. More than 50% of CVBR-infected MS-Z 293s responded to tebufenozide, however, exceeding all DE-based vectors with murA. More dramatic differences were seen in CV-1 and FF12 MS-Z infected cells. CVDE alone was not inducible in either cell type, indicating that the endogenous level of RXR in either CV-1s or primary fibroblasts is not sufficient to support VDE-mediated transactivation. When the endogenous level of RXR is supplemented, as in CVDEiR, 3-10% of the infected population was observed to strongly respond to murA. CVBR, however, was clearly the most potent transactivator in the CV-1 and FF12 cell types with 20% of CV-1 and 55% of CVBR-infected FF12 cells histochemically positive.

Luciferase activity in MS-Luc cells infected in parallel with the same receptor viruses confirmed that CVBR was not the most potent transactivator in the CV-1 and FF12 cell types and that all receptor types were functional in 293 cells. The MS-Z data suggest that the low absolute level of CVDEiR induction in CV-1 and FF12 cells with murA most likely reflects a small number of robustly responding cells in a largely nonresponsive population as opposed to low level induction by most MS-luc-infected cells.

Example 18
Transient transfections

Transient transfections are performed by calcium-phosphate coprecipitation employing standard methods (see Sambrook et al., in Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, New York, New York (1989)). All tissue culture experiments are performed using DMEM 10% FBS in a 10% CO_2 incubator unless otherwise specified. All transfections are performed in triplicate in 24-well Costar plates using CV-1 cells at an approximately $5X10^4$ plating density. Immediately following transfection, 1 μm ligand in 20% EtOH/PBS is added to wells. After 40 additional hours of incubation, cells are harvested and luciferase activity measured in an analytical bioluminescence photometer. Cell extracts are simultaneously examined for LacZ activity of the internal control by standard methods. Bar graph levels represent relative luciferase activity after correction using internal control values. All molecular biology enzymes and reagents used in this study are provided by either NEB (Beverly, MA) or Strategene (La Jolla, CA).

Example 19
Prolonged transgene expression for *ex vivo* gene therapy applications

Autologous, explanted skin fibroblasts genetically modified to express tyrosine hydroxylase (TH), the enzyme responsible for synthesis of L-dopa and the precursor for the neurotransmitter dopamine, have proven to ameliorate a loss of local dopamine in animal models of Parkinson's disease. Although quite effective in providing dopamine to the area of neural damage, this technique is only therapeutically useful for 2-3 weeks following transplantation. The decrease in effectiveness has been traced back to a dramatic loss of TH transgene expression in transplanted, post-mitotic cells. The use of invention constructs allows one to overcome the loss of transgene expression by providing stimulation of the retroviral LTR promoter through either ligand-activated transactivating complexes or through constitutive transactivating receptor variants. In this way, transgene expression may be maintained for longer periods of time, even indefinitely if desired.

Stem-type cells, such as stem cells of the hematopoetic system, nervous system, or embryo, could be infected or transfected with VbR regulated transgenes and subsequently implanted into adult, fetus, or early embryonic animals for either therapeutic or research purposes. Cells of the hematopoetic system could conditionally express proteins producing blood clotting factors such as factor IX, metabolic factors such as glucocerebrosidase, or protective factors including anti-HIV proteins.

Example 20
Prolonged transgene expression for *in vivo* gene therapy applications

The small size and regulatory capacity of VbRs lend themselves to use in recombinant retroviruses as a method of gene transfer. VbRs have been introduced into both MLV and lentiviral-based retroviral systems. When these viruses are introduced directly into target cells of either a mature or developing organism, expression of the virally encoded transgenes may be regulated by systemic addition of ligands such as tebufenozide and derivatives. An example of a disease that could be theoretically ameliorated by application of *in vivo* VbR encoding retroviruses is Parkinson's disease, described above.

Another application would be to use VbR encoding retroviruses as an anti-viral agent. Lentiviral vectors with regulated properties and harboring suicide genes or "protective" proteins could be used as a means of conditionally depleting or destroying HIV-positive cells.

Example 21
Modulated transgene expression for either *in vivo* or *ex vivo* gene therapy application

Treatment of Parkinson's disease with the chemical precursor of dopamine, L-dopa, has proven effective in ameliorating many of the deficits of Parkinsonism. With time, however, patients become refractory to L-dopa therapy, with the deleterious effects of chronic treatment outweighing even the serious symptoms of the disease itself. Eventually, patients are left with few therapeutic options. While the transplantation of TH expressing cells may be effective when constantly producing low-levels of L-

dopa, a potentially far more beneficial approach would be to allow the physician some degree of control over L-dopa production in the patient. This would allow sufficient control to ensure that the transgenic factor is expressed at appropriate therapeutic levels. At times when endogenous systems are capable of providing full function, the transgene may be allowed to become quiescent and transcriptionally inactive until needed again. Because the transcriptional induction of the invention retroviral constructs is dependent on an exogenous ligand, expression of an integrated therapeutic transgene can be placed under the control of the physician and patient.

Example 22

Construction of a BORIS vector expressing PolyQ-GFP

PolyQ-GFP are constructed from an existing GFP clone, GFP-SKSP. PCR primers encoding a novel restriction site and ATG start codon at the 5' end, and an Age I site in-frame with the GFP ORF at the 3' end are obtained for PCR amplification of huntingtin polyQ regions. These PCR products are inserted in-frame into GFP-SKSP to produce the fusion constructs of pQ-GFP. The clones are analyzed for mutations and function by sequencing and in vitro translation to confirm the creation of a pQ-GFP ORF. Alternate antibiotic resistance BORIS regulatable retroviral vectors are produced by removing the L-Histidinol gene from plasmid LSHL, by PCR, and inserting it in place on the *neo*^{r} into a BORIS vector for production of BORIS-LHis. The plasmid are large-scale prepped. A preliminary test for function is transient transfection experiments into transfectable cell types such as 293 cells and examination for the expression of beta-galactosidase or GFP reporter genes. Retrovirus are transiently produced and used for infection of target cells. Infected cells are selected in L-Histidinol containing media to ensure proper

selection. polyQ-GFP is then inserted into regulated retroviral vectors. pQ-GFPs are excised from SKSP and inserted into the regulatable BORIS and BORIS-LHis vectors. The plasmids are large-scale prepped. Function is determined by examination for green fluorescence and altered phenotype. Additional plasmid are used for transient production of retroviruses.

Primary rat fibroblasts, and the *hcnlv-myc* are selected for initial testing of BPQGS. Both are infected with the appropriate vector(s) place under selection, and individual resulting colonies subcloned. The subclones are expanded and passaged for analysis. Infected cells are treated or not with their respective ligands and the resulting phenotypes analyzed. Preliminary analysis will examine GFP fluorescence and distribution within the cells. At timed intervals following ligand induced expression, the pattern of fluorescence are recorded to determine if the fusion protein changes in distribution. Morphological characteristics of stimulated cells their growth rates, and cell death will also be examined at intervals following induction. For the *hcn/v-myc* cells, the consequences of polyQ-GFP expression in the presence and absence of tetracycline are compared.

Example 23

Aggregation of polyQ-fusion proteins

GFP fusion proteins with long (97Q), short (13Q), or no N-terminal polyQ tracts were constructed. Unique restriction endonuclease sites were designed into the regions flanking the polyQ tracts to allow insertion of an SV40 nuclear localization signal (NLS) or other sequences either C-terminal to the polyQ tract. Transient transfection analysis of these constructs in the NIT expression vector revealed that cells transduced with GFP

only or 13Q-GFP expressed GFP uniformly throughout their cytoplasm and that the NLS variant 13QN-GFP localized predominantly in the nucleus with a uniform distribution. 97Q- and 97QN-GFP constructs, on the other hand, initially manifested as uniform GFP positive fluorescence throughout the cytoplasm or nucleus, but within 24-72 hours after transfection condensed into bright granules indicative of intracellular aggregates (IAs). IAs were localized within the cytoplasm, surrounding the nuclear envelope, and within the nucleus itself (Figure 1B). The presence of the SV40 nuclear localization signal (NLS), though observed to efficiently direct localization of either 13Q- or 97Q-GFP variants to the nucleus, resulted in a quantifiable difference in the rate of IA formation after transfection.

Example 24

GFP-fusion proteins with short polyglutamine tracts are recruited into IAs

It was sought to determine whether overexpression of 13Q-GFP constructs would inhibit or interfere with the formation of IAs from 97Q-fusion constructs. 13Q-GFP was not observed to exert any inhibitory influence on 97Q-GFP mediated IA formation, and contrary to expectations, appeared to result in a qualitative difference in the IAs that formed. To test whether the 13Q-GFP fusion proteins were participating in the formation of IAs, the GFP-encoding portion of 97Q-GFP (97QΔ) was deleted and this construct was cotransfected with 13Q-GFP into cultured 293 cells. Fluorescent IAs were readily observed in a large population of the GFP-positive cells within the 24-72 hr. time course observed for 97Q-GFP alone. 13Q-GFP/97QΔ doubly infected cells were distinctive from 97Q-GFP cells because they continued to display faint uniform cytoplasmic GFP positivity in addition to the formation of predominantly nuclear IAs, whereas IA formation within

97Q-GFP cells generally results in condensation of all GFP fluorescence into IAs leaving other areas of the cell nonfluorescent and dark. Control cotransfection of NIT-97QΔ with NIT-GFP constructs with no N-terminal polyglutamine tracts did not result in the formation of IAs under any conditions tested, indicating that the short polyQ tract of 13Q-GFP is necessary for participation in IA formation.

Preliminary studies of 13Q-GFP recruitment were performed with a molar excess of 97QΔ plasmid. To further determine the molar ratio of 97Q to 13Q required for participation of 13Q in IA formation, 13Q-GFP was titrated with 97QΔ in a cotransfection assay and the percent of IA containing cells quantified relative to the total number of visibly GFP positive cells. No IAs were observed with non cotransfected 97QΔ or at a 0.1 molar ratio of 97QΔ plasmid; however, at a 0.5 molar ratio of 97QΔ to 13Q-GFP, 10% of GFP positive cells contained GFP positive IAs. At a 1:1 ratio, nearly 40% of GFP positive cells also displayed IAs. At 10:1 ratio of NIT-97QΔ to NIT-13Q-GFP, over 75% of positive cells displayed IAs 60 hours post-transfection.

Example 25

Characterization of IAs formed from a combination of long and short polyQ-reporters

To better characterize the formation of IAs in the presence of both long and short polyQ tracts, 97QΔ was modified to contain an in-frame fusion of the LacZ gene encoding β-galactosidase (97QZ). Preliminary transient transfection studies of 97QZ revealed that like 97Q-GFP, 97QZ condensed within transfected cells to form both immunopositive and histochemically reactive IAs. NIT-97QZ and NIT-13Q-GFP were cotransfected into cells and 60 hours later, performed immunohistochemistry and histochemical

reactions on the transfected cells to simultaneously examine both reporter constructs. Although the β-galactosidase (β-gal) histochemical reaction obscures GFP fluorescence, through slight modification of standard staining techniques and limiting the time of exposure to reagents, a low level of GFP fluorescence could be preserved and imaged along with the characteristic blue stain of the β-galactosidase reaction. A population of IAs are observed that are both GFP and β-galactosidase positive. Numerous large β-gal only positive IAs are also present, presumably from the aggregation of 97QZ only (or predominantly). Curiously, several IAs that appear only positive for GFP were also observed using the histochemical stain and native GFP fluorescence imaging.

Further exploration of this observation was performed using confocal microscopy and immunohistochemistry to identify reporters within IAs. As in the above study, individual IAs were clearly observed to be positive for both reporters while other IAs were positive only for β-gal for GFP. Taken together, these results clearly indicate that long polyglutamine tract-reporter fusion constructs form IAs *in vitro*, and that the forming aggregates can either recruit short polyglutamine-reporters into the aggregate, induce them to aggregate on their own, or both.

Example 26

Construction of cell lines with regulated IA formation and 97Q-GFP expression

The study of IA formation has thus far relied heavily on transient transfection analysis and has been hampered by a lack of stable IA-forming cultured cell populations. Through the use of retroviral vectors, a variety of stably cultured cell lines conditionally or chronically expressing the polyQ-GFP constructs were constructed. Two regulatable

vector systems, an induced expression system based on a modified ecdysone receptor variant (VbR) and a repressed expression system based on the tetracycline transactivator (TTA), were used to create cell lines with inducible IA formation. These two vectors are best suited to different ranges of induction in cultured cell types. VBR-based B3 vectors are particularly suited to applications where very low basal expression levels that can be rapidly stimulated to moderately high levels are desired. The TTA-based NIT vector tends not have the broad-based low level of "uninduced" expression characteristic of the BORIS vectors (identified in SEQ ID NO:6); however, because transactivator expression and transactivation is continuous in the absence of ligand, very high transgene expression levels may be achieved through the use of this vector.

97Q-GFP was cloned via a custom shuttle vector into both retroviral vectors. Transiently produced B3-polyQ-GFP retroviruses (B3-13Q or B3-97Q) and TTA-based NIT polyQ-GFP retroviruses (NIT-13Q or NIT-97Q) were used at low M0I to infect several well-characterized cell lines used previously in transduction studies of IA formation. COS-7 cells were selected because several studies had examined IA formation and IA-induced apoptosis in this cell type. 293 cells were utilized for the same reason and also because they are amenable to efficient transfection and would have significant utility in the development of gene-based factors that interact with or interrupt IA formation. Primary rat fibroblasts were selected because they are non-transformed and readily enter a prolonged post-mitotic state following contact inhibition at high culture densities.

Pilot studies with reporter constructs revealed that the B3 vector functioned most efficiently in the 293 and COS cell lines and that NIT functioned at highest

efficiency in the primary fibroblast line. Individual cell lines infected with the appropriate vector were placed under G418 selection to remove uninfected cells. For B3-polyQ infected 293 and COS-7 cells, the resulting resistant population was pooled and passaged without ligand. For NIT-polyQ infected primary fibroblasts, cells were cultured with the tetracycline analog doxycycline prior to, during, and after infection to prevent chronic polyQ overexpression to disrupt growth of resistant colonies. 293 and COS-7 cells were also infected with NIT constructs and cultured without added doxycycline to determine the effects of chronic overexpression of 97Q-GFP on these cell types as described below.

Example 27

Stable IA-forming 293 cells

G418-resistant B3-97Q 293s displayed little to no visible GFP fluorescence during culture in the absence of ligand; however, introduction of 1 μM tebufenozide into the culture medium resulted in a rapid increase of cytoplasmic fluorescence within 24 hours post-induction. Nearly 100% of cells would display GFP-positive fluoresence within 24-72 hours of induction. After 24 hours, only widely scattered cells would contain small, detectable IAs. By 48 hours, >8% of the cell population would contain one or more IAs increasing to >15% of the population between 48 and 72 hours. Occasional IA containing cells would be observed floating in the culture medium; however, there was no evident "wave" of cell death at the 72 hour time point or later times. Passage of B3-97Q 293s in the presence of ligand for three passages resulted in only a small but detectable decrease in the percentage of cells containing IAs (15% to 11%).

NIT-97Q infection sheds further light on the toxic potential of long-polyglutamine repeats on 293 cells. Surprisingly, colonies of 100% IA-positive 293 cells survived for over two weeks and apparently divided and grew, in some cases, to colonies of hundreds of cells (Figure 4D). By the end of 21 days of culture; however, the majority of IA positive colonies had disappeared leaving only scattered IA positive cells and numerous colonies of GFP positive cells without IAs that eventually grew to fill the plates. The expression time-course and level of B3- AND NIT-13Q cells paralleled the 97Q cells although no IAs were formed.

Example 28
Stable IA-forming COS-7 cells

G418-resistant B3-97Q COS-7s displayed no visible GFP fluoresence during culture in the absence of ligand in a manner almost identical to the 293 B3-97Q population. Addition of 1 μM tebufenozide into the culture medium resulted in a rapid increase of cytoplasmic fluorescence within 24-48 hours post-induction. IA formation peaked at 4-5% of total cells (at any given moment) 72-96 hours post-induction. Unlike the B3-97Q 293 cells, a continuous population of G418-resistant cells did not display even GFP fluoresence suggesting that one or more components of the regulatable system were lost from the COS population. Like B3-97Q 293 cells, IA containing COS-7 cells were also observed floating in the culture medium; however a depletion of the expressing population was not observed presumably because the continuous formation of IAs in other cells.

The effects of chronic 97Q-GFP overexpression in infected COS-7 cells was examined using the NIT vector. Like NIT-97Q infected 293s, colonies in which the majority

of cells contained large and numerous IAs were readily observed. Unlike the overexpressing 293 cells; however, several selected colonies continued to grow and express IAs in a continual percentage of the population even after passage. Three lines, COSN97Q-4, 6 and 7, continually express IAs after continuous growth for over four weeks. COSN97Q-7 displays the highest level of spontaneous IA formation. At any given time, approximately 6% of the total population contains readily observable IAs. CV-1 cells, a predecessor of the COS-7 cells were also infected with NIT-97Q and formed IA containing colonies in much the same manner as COS cell populations.

Example 29
Stable IA forming primary fibroblasts

B3-97Q infected primary fibroblasts (FF12s) became faintly GFP fluorescent within 72 hours of tebufenozide treatment but did not achieve a maximum level of 97Q-GFP expression sufficient to result in IA formation in a significant population of cells. The NIT vectors were employed to generate regulated 97Q-GFP expression. Infected NIT-97Q FF12s were infected and cultured continuously in the presence of doxycycline (dox) to prevent IA formation. Twelve individual colonies of NIT-97Q FF12s were selected in order to screen for individual cell populations with low dox inhibited expression and high 97Q expression in the absence of ligand. One line, FFN97Q-5 displayed GFP fluoresence within 90% of the confluent cell population within 5 days of ligand removal. IAs began forming on day 5 and continued to increase through day 8. Approximately 15% of the total cell population displayed IAs 10 days post ligand removal. IA formation in the confluent FFN97Q-5 population was distinctive from other cell types examined in that it was occasionally exclusively nuclear and at other times

predominantly cytoplasmic in an unusual pattern within the cell body. IAs with a distinctive stellate or tri-partite appearance were often observed. Large cytoplasmic aggregates of multiple star-like IAs were also occasionally seen. Curiously, confluent FFN97Q-5 cells with multiple cytoplasmic or nuclear IAs appeared visually normal. Floating dead cells with or without IAs were rarely observed and in numbers no different from control wells. FF970Q-%s cultured in the presence of dox neither expressed visible 97Q-GFP nor formed IAs.

Example 30
Use of invention for efficient production of transgenic animals

Transgenic animals are generally produced by either pronuclear injection of DNA or by transfection of embryonic stem (ES) cells followed by selection and injection of the stem cell into the inner cell mass of very early embryos. Pronuclear injection results in approximately 5-10% stable gene transfer in the production of transgenic mice. The use of ES cells in producing transgenics is likewise inefficient in generating mosaics with germ-line transmission of the transgene. It was proposed in the mid-1980's to use retroviruses to transfer transgenes with high efficiency into early embryos or ES cells to dramatically enhance the odds of producing transgenic animals. All attempts at this failed, not because the virus was incapable of stably integrating into the target cell genome, but because the integrated provirus did not express any of the genes encoded within the viral transcriptional cassette.

The present invention is capable of overcoming the transcriptional block to result in germ-line transgenic animals with full expression from the integrated transgene.

In addition to expressing the transgene, the level of transcription may still be regulated by controlling the supply of ligand to the transgenic animal. The increased efficiency of producing transgenic animals by retroviral infection should open up the way to producing mutant animals of a variety of species previously impractical for genetic modification because of the potential cost of producing a large number of non-positive animals by classical methods.

To produce transgenic mice, the following nucleic acid constructs are prepared and subsequently injected into fertilized eggs: CD3-VbR and a ligand inducible ß-gal reporter. Two separate lines of transgenic mice are generated harboring either a ligand inducible reporter, or a T-cell specific expression construct of VbR, respectively. The former are referred to as reporter mice, the latter are referred to as receptor mice, and double transgenic mice are referred to as receptor/reporter mice. Constructs CD3-VbR are injected, while the reporter is injected alone. Primary genotyping is performed by Southern blot analysis and the transmission of transgenic mice is monitored by dot blot analysis. Receptor mice are analyzed for VbR expression by Northern blot analysis of RNA collected from these mice. For Northern blot analysis, 15μg of total RNA obtained from the thymus, and various tissues as a control, is run on a denaturing gel and blotted onto a nitrocellulose membrane. The blot is probed with a radiolabeled ß-gal-specific probe and exposed on film for 2 days. In addition, the transgene can be transferred to the offspring as expected by Mendelian genetics.

Example 31
Use of retroviral constructs in the invention for efficient gene transfer to developing embryos

Since the present invention can effectively overcome the block of viral expression in embryonic cells, invention constructs are a potent tool in the delivery of transgenes to somatic tissues of a developing embryo. With many diseases, considerable damage is done during embryonic development so that therapies applied after birth are essentially ineffective to ameliorate the disease phenotype.

The present invention provides methodology where one can infect cells of the embryo and can provide therapeutic factors to the developing fetus either constitutively, or under the regulation of exogenously produced ligand.

Example 32
Use of vector constructs in invention with inducible high titers

One obstacle to the use of retroviruses as gene transfer agents is that titers of retroviruses from existing producer cell lines are only on the order of $1X10^4$ or $1X10^5$. By using a retroviral construct of the invention having intact enhancers and regulatory elements, expression of the retrovirus may be induced by greater than ten-fold, resulting in correspondingly higher titers of infectious virus.

While the invention has been described in detail with reference to certain preferred embodiments thereof, it will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

That which is claimed is:

1. A nucleic acid construct encoding a receptor peptide comprising:

(i) a ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily

(ii) an activation domain, and

(iii) a DNA binding domain;

wherein said receptor peptide activates regulatory element(s) in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

2. A nucleic acid construct according to claim 1 wherein said ligand for said ligand binding domain is a non-steroidal compound not normally present in a mammalian cell.

3. A nucleic acid construct according to claim 1, wherein said ligand binding domain and hinge region are derived from a lepidopteran species.

4. A nucleic acid construct according to claim 1, wherein said ligand binding domain and hinge region are derived from *bombyx mori*.

5. A nucleic acid construct according to claim 1 wherein said ligand binding domain has substantial sequence identity with amino acid sequence 393-508, as set forth in SEQ ID NO:3.

6. A nucleic acid construct according to claim 1 wherein said ligand binding domain has substantial

sequence identity with amino acid sequence 393-586, as set forth in SEQ ID NO:3.

7. A nucleic acid construct according to claim 1 wherein said hinge region has substantial sequence identity with amino acid sequence 273-362, as set forth in SEQ ID NO:3.

8. A nucleic acid construct according to claim 1 wherein said hinge region has substantial sequence identity with amino acid sequence 273-362, as set forth in SEQ ID NO:3.

9. A nucleic acid construct according to claim 1 wherein said DNA binding domain is obtained from a DNA-binding protein.

10. A nucleic acid construct according to claim 9 wherein said DNA binding domain is obtained from a member of the nuclear receptor superfamily.

11. A nucleic acid construct according to claim 10 wherein said DNA binding domain is non-mammalian.

12. A nucleic acid construct according to claim 1 wherein said receptor peptide has substantial sequence identity with the amino acid sequence set forth in SEQ ID NO 5.

13. A nucleic acid construct according to claim 1, wherein said nucleic acid construct further comprises a regulatory element operatively associated with exogenous nucleic acid(s).

14. A nucleic acid construct according to claim 13 wherein said regulatory element comprises a hormone response element.

15. A nucleic acid construct according to claim 14 wherein said construct comprises a plurality of hormone response elements.

16. A nucleic acid construct according to claim 13 wherein said exogenous nucleic acid(s) is (are) selected from wild type and therapeutic nucleic acid(s).

17. A nucleic acid construct according to claim 16 wherein said wild type nucleic acid(s) is (are) selected from genes which encode products:

the substantial absence of which leads to the occurrence of a non-normal state in said host; or

a substantial excess of which leads to the occurrence of a non-normal state in said host.

18. A nucleic acid construct according to claim 17, wherein said exogenous nucleic acid is a CAG repeat encoding a polyglutamine polypeptide.

19. A nucleic acid construct according to claim 16 wherein said therapeutic nucleic acid(s) is (are) selected from those which encode products:

which are toxic to the cells in which they are expressed; or

which impart a beneficial property to said host.

20. A nucleic acid construct according to claim 19 wherein said therapeutic nucleic acid(s) modulates the

transcription and/or translation of endogenous nucleic acid(s).

21. A nucleic acid construct according to claim 1, further comprising nucleic acid encoding a selectable marker.

22. A nucleic acid construct according to claim 21 wherein said selectable marker is a radiolabeled molecule, a fluorescent molecule, an enzyme, or a ligand.

23. A nucleic acid construct according to claim 22 wherein said enzyme is an antibiotic resistance gene or a gene which enables cells to process a metabolic intermediary.

24. An expression cassette comprising a promoter operatively associated with a nucleic acid construct according to claim 1.

25. An expression cassette according to claim 24, wherein said promoter is selected from inducible, constitutive and/or tissue specific promoters.

26. A viral vector capable of ligand-mediated expression, wherein said vector comprises an expression cassette according to claim 24.

27. A viral vector according to claim 26, wherein said viral vector further comprises a retroviral psi (Ψ) packaging signal and a 5' and/or a 3' long terminal repeat (LTR).

28. An animal cell in culture which is transformed with a nucleic acid construct according to claim 1.

29. A cell according to claim 28, wherein said cell is further transformed with a reporter vector comprising a regulatory element operatively associated with exogenous nucleic acid(s).

30. An animal cell in culture which is transformed with a nucleic acid construct according to claim 13.

31. An animal cell according to claim 30, wherein said cell is an embryonic stem cell.

32. A transgenic animal, wherein cells of said animal is transfected with a nucleic acid construct according to claim 1.

33. A transgenic animal, wherein cells of said animal is transfected with a nucleic acid construct according to claim 13.

34. A polypeptide receptor encoded by a nucleic acid construct according to claim 1.

35. A nucleic acid construct encoding a receptor peptide comprising :

(i) a ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily, and

(ii) a DNA binding domain, wherein said DNA binding domain is not derived from said non-mammalian member;

wherein said receptor peptide activates regulatory

element(s) in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

36. A nucleic acid construct according to claim 35, wherein said receptor peptide further comprises:

(iii) an activation domain.

37. A nucleic acid construct according to claim 35, wherein said nucleic acid construct further comprises a regulatory element operatively associated with exogenous nucleic acid(s).

38. An expression cassette comprising a promoter operatively associated with a nucleic acid construct according to claim 35.

39. A viral vector capable of ligand-mediated expression, wherein said vector comprises an expression cassette according to claim 38.

40. An animal cell in culture which is transformed with a nucleic acid construct according to claim 35.

41. A cell according to claim 40, wherein said cell is further transformed with a regulatory element operatively associated with exogenous nucleic acid(s).

42. A polypeptide receptor comprising:

a VP16 activation domain,

a DNA binding domain, and

bombyx mori-derived ligand binding domain and hinge region.

43. A nucleic acid construct comprising a regulatory element derived from a bombyx receptor response element operatively associated with exogenous nucleic acid(s).

44. An animal cell in culture which is transformed with a nucleic acid construct according to claim 43.

45. An animal cell according to claim 44, wherein a polypeptide receptor is further introduced into said cell, said receptor comprising:

an activation domain,

a DNA binding domain,

and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily,

wherein said receptor peptide activates regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

46. A transgenic nonhuman mammal, wherein cells of said mammal are transfected with the nucleic acid construct according to claim 43.

47. A method for modulating the transcription of exogenous nucleic acid(s) in a host containing: ,

(i) a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element; and

(ii) a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge

region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain,

said method comprising administering to said host an amount of ligand effective to modulate the transcription of said exogenous nucleic acid(s); wherein ligand is not normally present in the cells of said host.

48. A method according to claim 47 wherein said receptor peptide is provided by a nucleic acid construct encoding said receptor peptide.

49. A method according to claim 48 wherein said receptor peptide is expressed under the control of a tissue specific promoter.

50. A method according to claim 47 wherein said exogenous nucleic acid(s) is (are) selected from wild type and therapeutic nucleic acid(s).

51. A method according to claim 50 wherein said wild type nucleic acid(s) is (are) selected from genes which encode products:

the substantial absence of which leads to the occurrence of a non-normal state in said host; or

a substantial excess of which leads to the occurrence of a non-normal state in said host.

52. A method according to claim 50 wherein said therapeutic nucleic acid(s) is (are) selected from those which encode products:

which are toxic to the cells in which they are expressed; or

which impart a beneficial property to said host.

53. A method according to claim 50 wherein said therapeutic nucleic acid(s) modulates the transcription and/or translation of an endogenous nucleic acid(s).

54. A method according to claim 47 wherein said regulatory element is not normally present in the cells of said host.

55. A method according to claim 47 wherein said DNA binding domain is obtained from a DNA-binding protein.

56. A method according to claim 55 wherein said DNA binding domain is obtained from a member of the nuclear receptor superfamily.

57. A method according to claim 55 wherein said DNA binding domain is a GAL4 DNA binding domain.

58. A method according to claim 47 wherein said receptor peptide further comprises an activation domain.

59. A method according to claim 47 wherein said host further contains a non-mammalian transactivator not normally present in the cells of said host and a transactivator responsive regulatory element not normally present in cells of said host,

wherein said transactivator responsive regulatory element controls transcription of said exogenous nucleic acid(s) or a second nucleic acid construct comprising a second exogenous nucleic acid(s).

60. A method according to claim 59 wherein said exogenous nucleic acid(s) is (are) under the further control of said transactivator responsive regulatory element.

61. A method according to claim 59 wherein said transactivator responsive regulatory element controls transcription of a second DNA construct comprising a second exogenous nucleic acid(s).

62. A method according to claim 59 wherein said transactivator responsive regulatory element comprises an operator which is responsive to said non-mammalian transactivator which confers responsiveness to antibiotics.

63. A method according to claim 59 wherein said transactivator is a ligand-mediated receptor which confers responsiveness to antibiotics.

64. A method according to claim 47 wherein said hinge region has substantial sequence identity with amino acid sequence 273-362, as set forth in SEQ ID NO:3.

65. A method according to claim 47 wherein said ligand binding domain has substantial sequence identity with amino acid sequence 393-508, as set forth in SEQ ID NO:3.

66. A method according to claim 47 wherein said ligand binding domain has substantial sequence identity with amino acid sequence 393-586, as set forth in SEQ ID NO:3.

67. A method according to claim 47 wherein said receptor peptide has substantial sequence identity with the sequence set forth in SEQ ID NO:3.

68. A method according to claim 47 wherein said receptor peptide has the sequence set forth in SEQ ID NO:3.

69. A method according to claim 47 wherein said ligand for said ligand binding domain is a non-steroidal compound.

70. A method according to claim 69 wherein said non-steroidal compound is a hydrazine.

71. A method according to claim 70 wherein said hydrazine is a 1,2-diacyl hydrazine, an N'-substituted-N,N'-disubstituted hydrazine, a dibenzoylalkyl cyanohydrazine, an N-substituted-N-alkyl-N,N-diaroyl hydrazine, an N-substituted-N-acyl-N-alkyl, carbonyl hydrazine, and an N-aroyl-N'-alkyl-N'-aroyl hydrazine.

72. A method according to claim 71 wherein said hydrazine is a diacyl hydrazine.

73. A method of inducing the transcription of an exogenous nucleic acid(s) in a host containing:

(i) a nucleic acid construct comprising a promoter

and said exogenous nucleic acid(s) under the control of a regulatory element;

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein expression of said receptor peptide is under the control of an inducible promoter, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain, and

(iii) said ligand for said ligand binding domain, wherein said ligand is not normally present in the cells of said host;

said method comprising subjecting said host to conditions suitable to induce expression of said receptor peptide.

74. A method of inducing the transcription of an exogenous gene in a host containing a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element; said method comprising introducing into a host:

a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear

receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain, and

a ligand for said ligand binding domain.

75. A method for the treatment of a host in need of gene therapy, said method comprising:

introducing into cells of said host:

(i) a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element;

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein expression of said receptor peptide is under the control of an inducible promoter, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain; and

administering, to said host, ligand for said ligand binding domain.

76. A method for the treatment of a host in need of gene therapy, said method comprising:

introducing into cells obtained from said host:

(i) a nucleic acid construct comprising a promoter and exogenous nucleic acid(s) under the control of a regulatory element; and

(ii) nucleic acid endocing a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain,

to provide modified cells;

reintroducing the modified cells into said host, and

administering, to said host, ligand for said ligand binding domain.

77. A method for producing transgenic animals capable of prolonged and regulated transcription of exogenous nucleic acid(s), said method comprising introducing into early-stage embryos or stem cells:

(i) a nucleic acid construct comprising a promoter and said exogenous nucleic acid(s) under the control of a regulatory element;

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor

superfamily which is not normally present in the cells of said transgenic animal, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

78. A method for the expression of recombinant products detrimental to a host organism, said method comprising:

transforming suitable host cells with:

(i) a nucleic acid construct comprising a promoter and exogenous nucleic acid(s) which express said recombinant product under the control of a regulatory element; wherein said regulatory element is not normally present in the cells of said host, and

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner for said ligand binding domain, and

growing said host cells to the desired level in the substantial absence of ligand for said receptor peptide; and

inducing expression of said recombinant product by

introducing into said host cells said ligand, which, in combination with said receptor peptide, binds to said regulatory element and activates transcription therefrom.

79. A method according to claim 78, wherein said exogenous nucleic acid is a CAG repeat which encodes a polyglutamine polypeptide.

80. A method for identifying compounds useful for treating neuodegenerative conditions, said method comprising:

introducing a test compound into a host cell transformed according to claim 79, and

identifying those test compounds that target intracellular aggregates of polyglutamines as compounds useful for treating neuodegenerative conditions.

81. A method for modulating the transcription of nucleic acid(s) in an *in vitro* system, said method comprising administering to said system an amount of said ligand effective to modulate the transcription of said nucleic acid(s); wherein said ligand is not normally present in said cellular system; wherein said system comprises:

(i) a nucleic acid construct comprising a promoter and said nucleic acid(s) under the control of a regulatory element; and

(ii) nucleic acid encoding a receptor peptide comprising a DNA binding domain, and the ligand binding domain and hinge region of a non-mammalian member of the nuclear receptor superfamily which is not normally present in

the cells of said host, wherein said receptor peptide activates said regulatory element in the absence of an exogenous dimer partner therefor and in the presence of a ligand for said ligand binding domain.

82. A nucleic acid construct comprising nucleic acid encoding a VP16 activation domain, a DNA binding domain and *bombyx mori*-derived ligand binding domain and hinge region.

83. A nucleic acid construct comprising a promoter and nucleic acid encoding a tetracycline-controlled transactivator, a VP16 activation domain, a DNA binding domain and *bombyx mori*-derived ligand binding domain and hinge region, in operative association with one another.

84. A vector comprising:

a first reporter under the control of a regulatory element,

a second reporter under the control of an operator which is responsive to a ligand-mediated receptor which confers responsiveness to antibiotics, and

a nucleic acid construct comprising a promoter operatively associated with nucleic acid encoding a tetracycline-controlled transactivator, a VP16 activation domain, a DNA binding domain and *bombyx mori*-derived ligand binding domain.

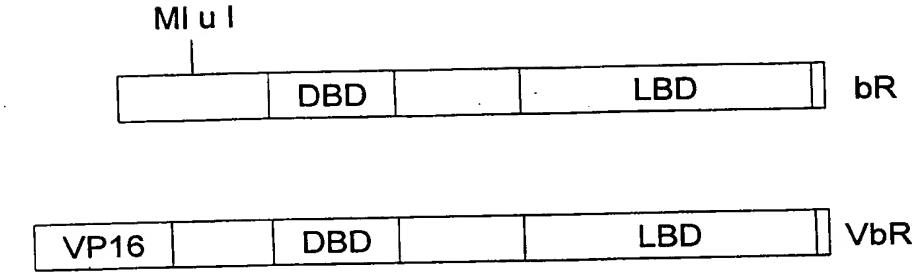


Figure 1A

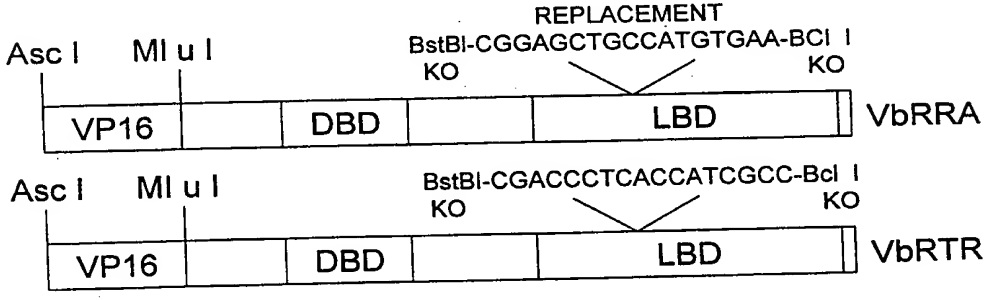


Figure 1B

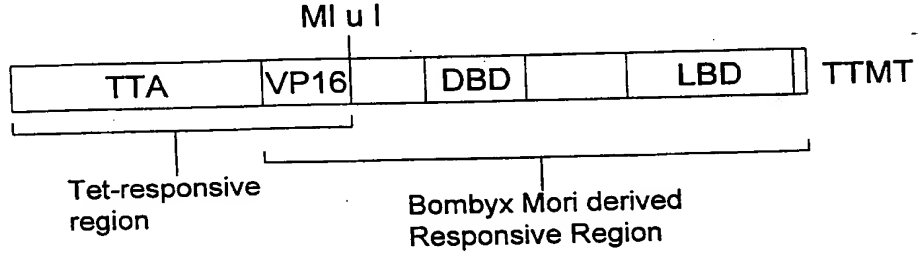


Figure 1C

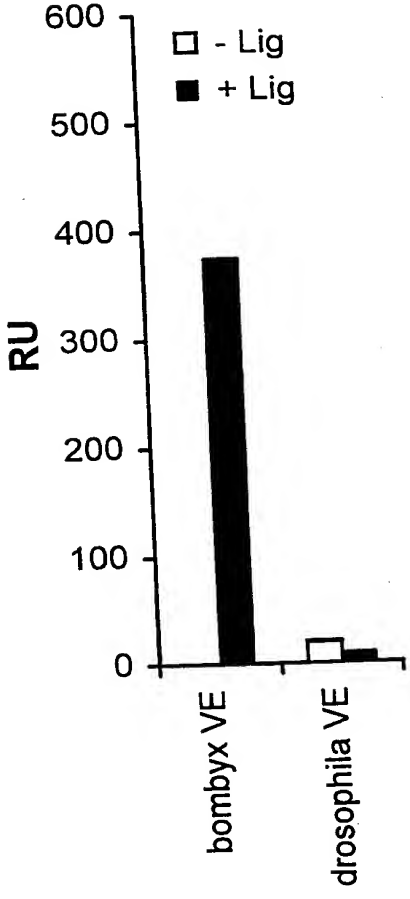


Figure 2A

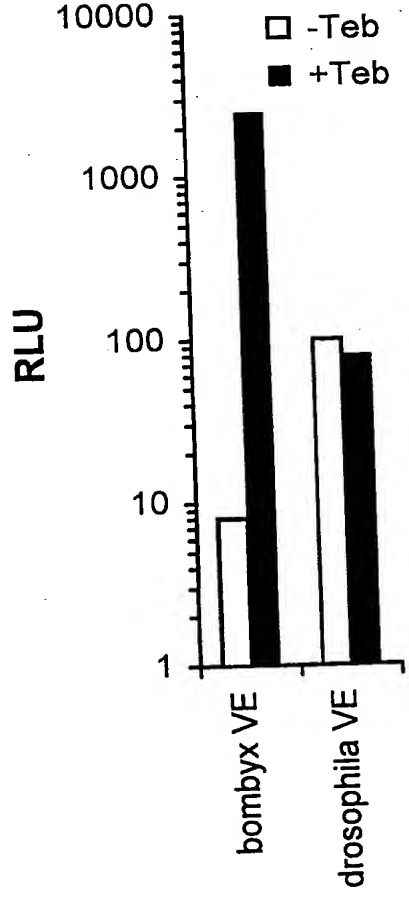


Figure 2B

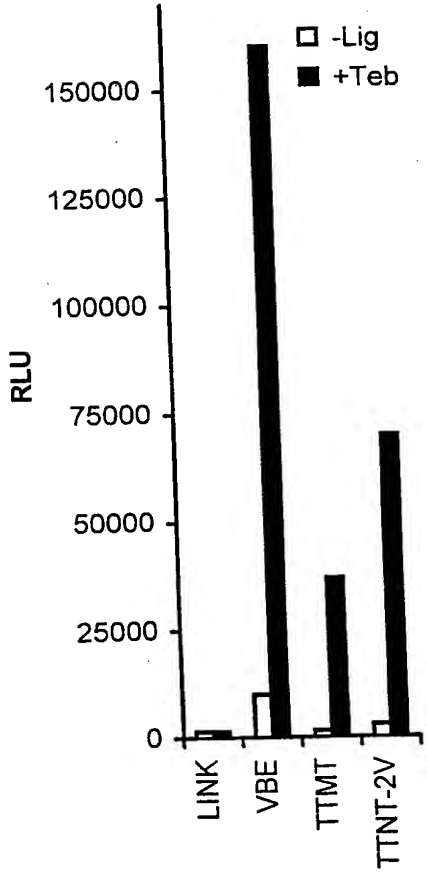
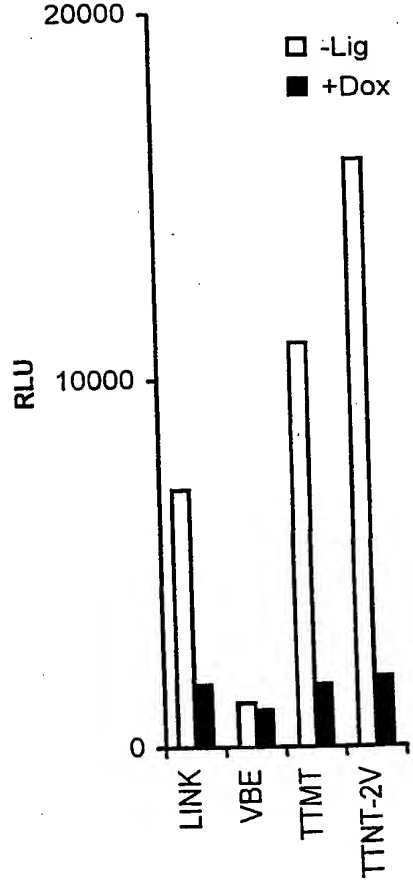


Figure 2C



Figur 2D

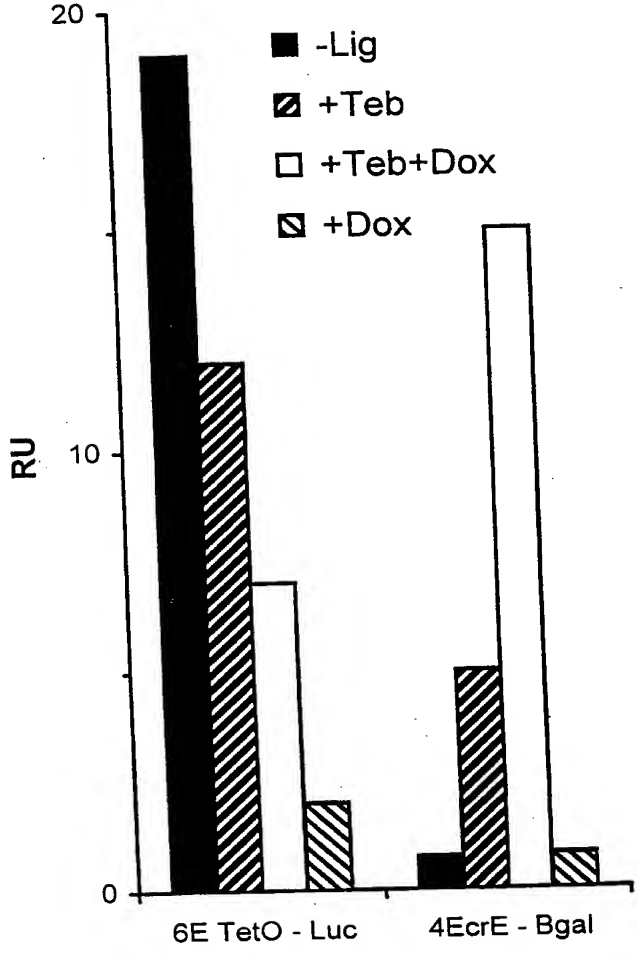
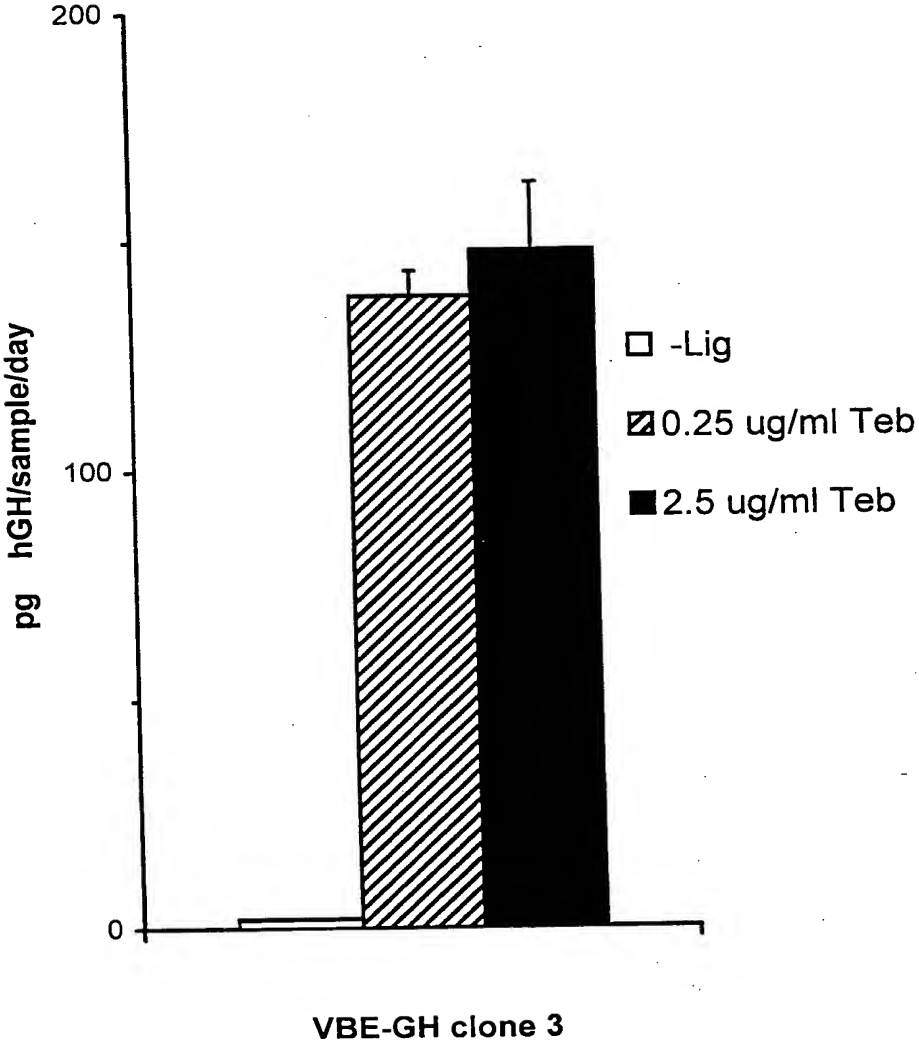


Figure 2E


200
100
0
pg hGH/sample/day
-Lig
0.25 ug/ml Teb
2.5 ug/ml Teb
VBE-GH clone 3

Figur 3

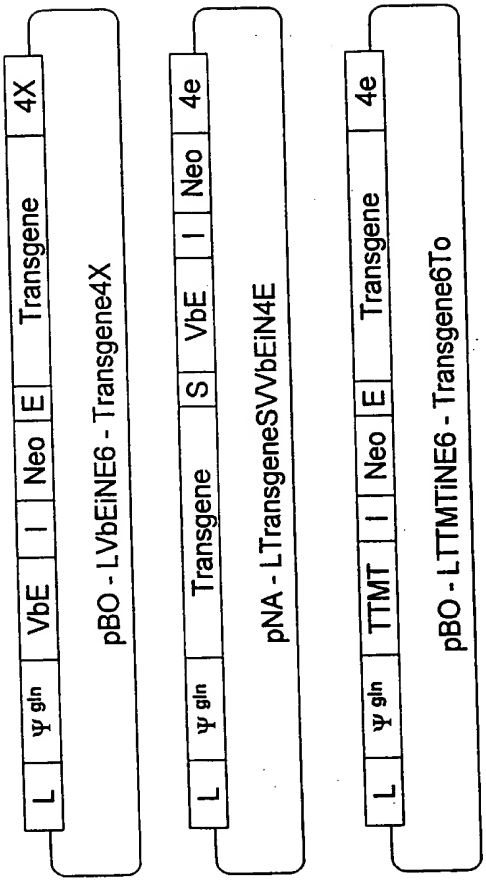

L
Ψ gln
VbE
I
Neo
E
Transgene
4X
pBO - LVbEiNE6 - Transgene4X
L
Ψ gln
Transgene
S
VbE
I
Neo
4e
pNA - LTransgeneSVVbEiN4E
L
Ψ gln
TTMT
I
Neo
E
Transgene
4e
pBO - LTTMTiNE6 - Transgene6To

Figure 4

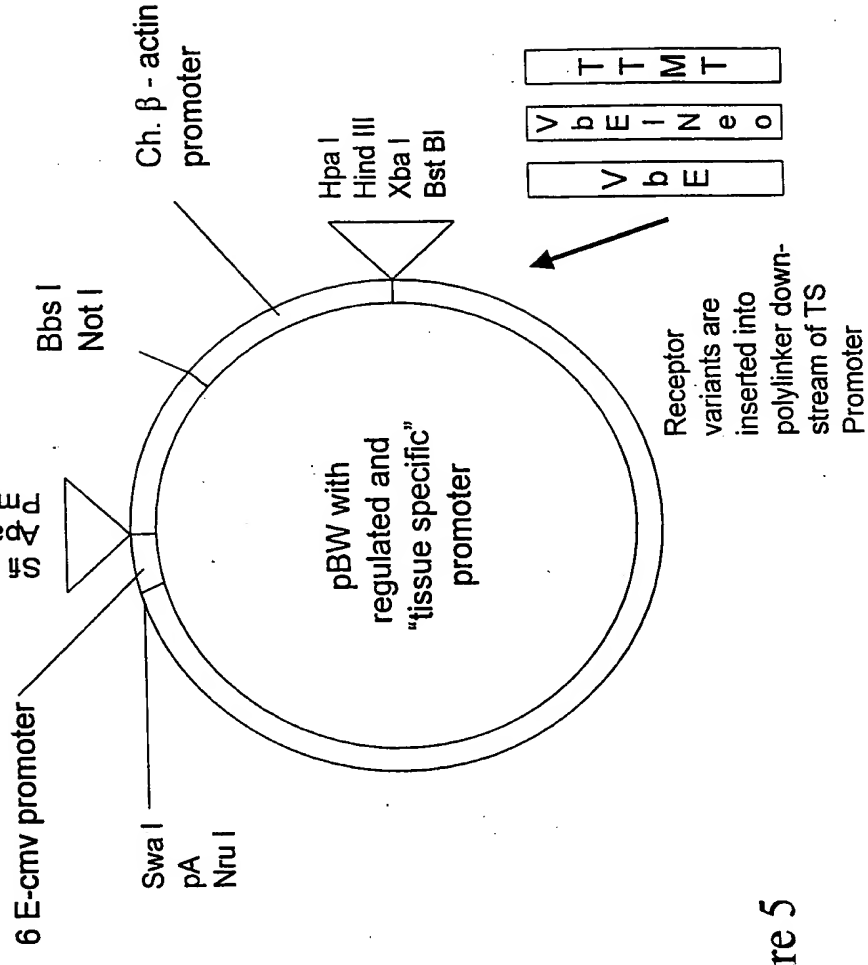

Ch. β - actin promoter
Hpa I
Hind III
Xba I
Bst BI
Bbs I
Not I
Sfi I
Pac I
Xho I
6 E-cmv promoter
Swa I
pA
Nru I
pBW with regulated and "tissue specific" promoter
Receptor variants are inserted into polylinker down-stream of TS Promoter

Figure 5

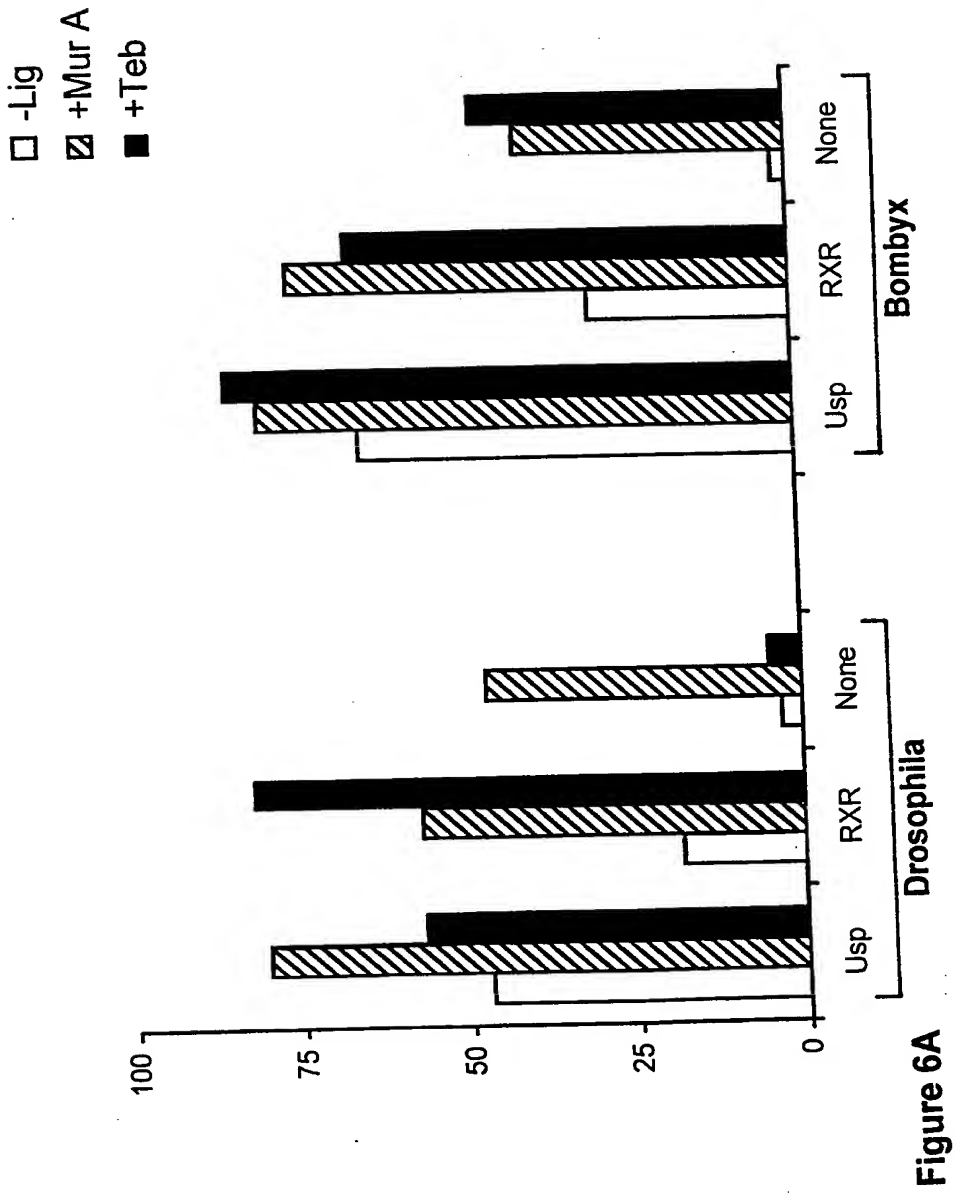

-Lig
+Mur A
+Teb
100
75
50
25
0
Usp
RXR
None
Drosophila
Usp
RXR
None
Bombyx

Figure 6A

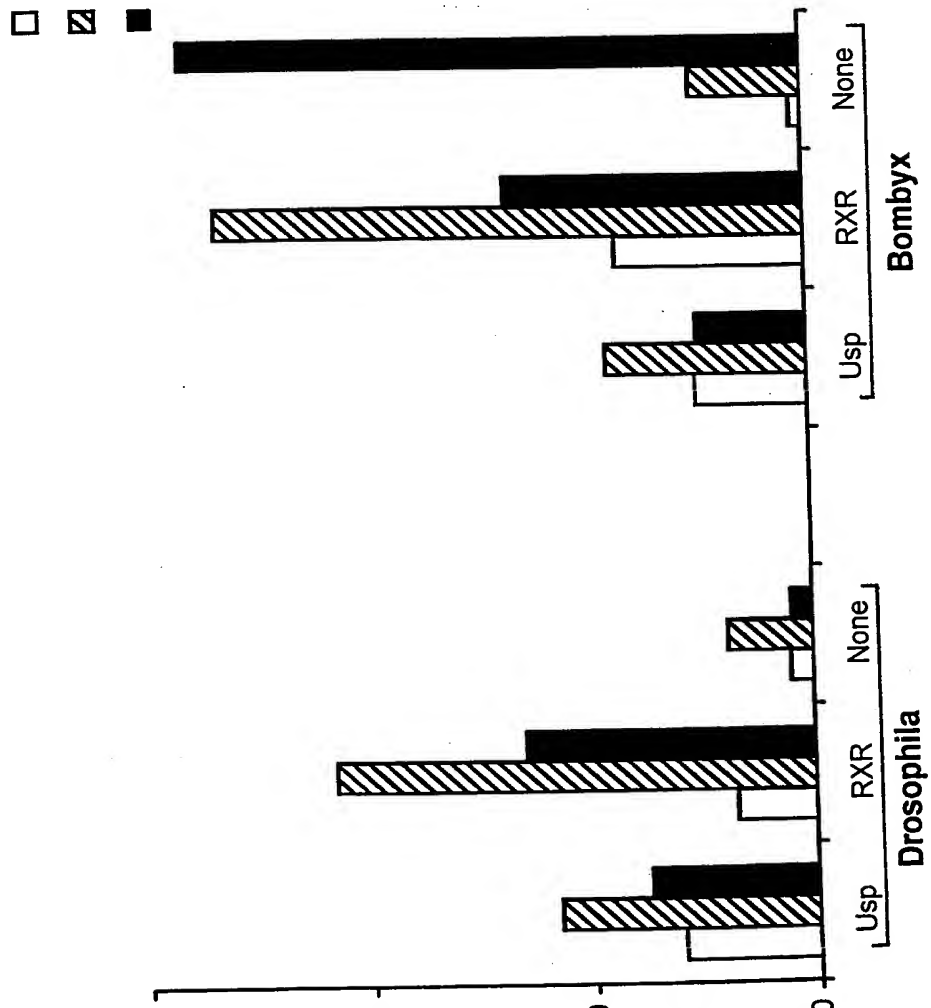

-Lig
+Mur A
+Teb
60
40
20
0
Usp
RXR
None
Drosophila
Usp
RXR
None
Bombyx

Figure 6B

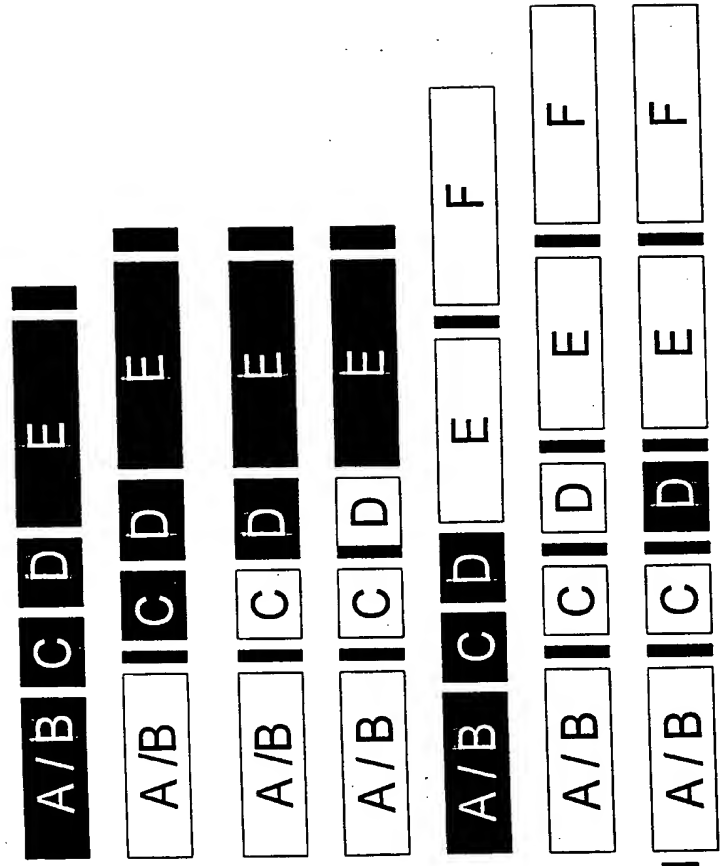

BE
A/B
C
D
E
DEBE- A
A/B
C
D
E
DEBE- C
A/B
C
D
E
DEBE- D
A/B
C
D
E
BEDE
A/B
C
D
E
F
DE
A/B
C
D
E
F
DEBH
A/B
C
D
E
F

Figure 7A

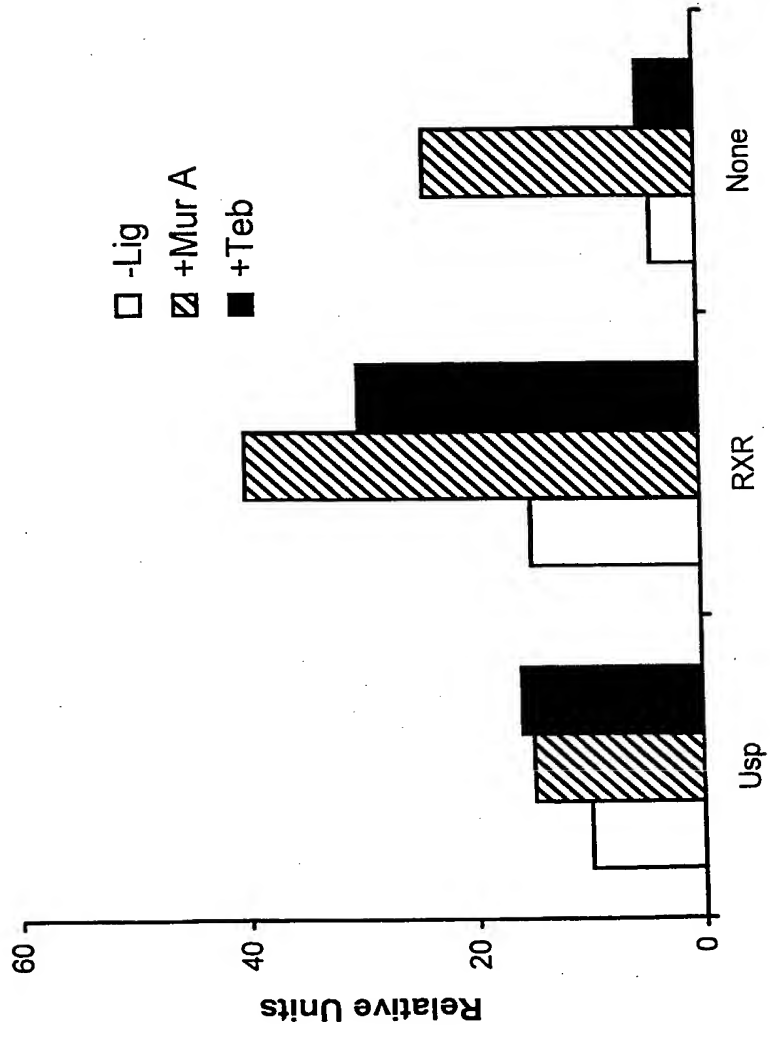


Figure 7B

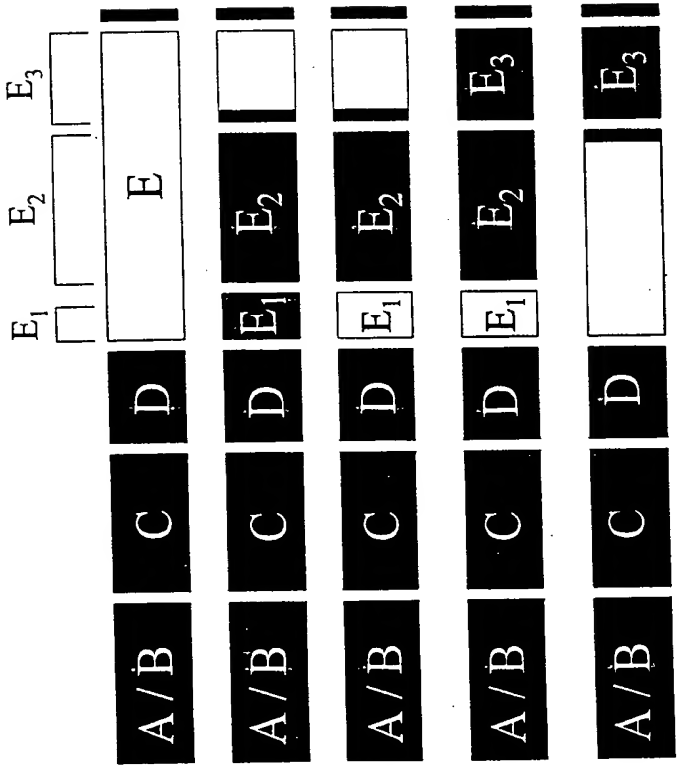

E1
E2
E3
BEDB
A / B
C
D
E
BKE
A / B
C
D
E1
E2
BAE
A / B
C
D
E1
E2
BAB
A / B
C
D
E1
E2
E3
BEB
A / B
C
D
E3

Figure 8

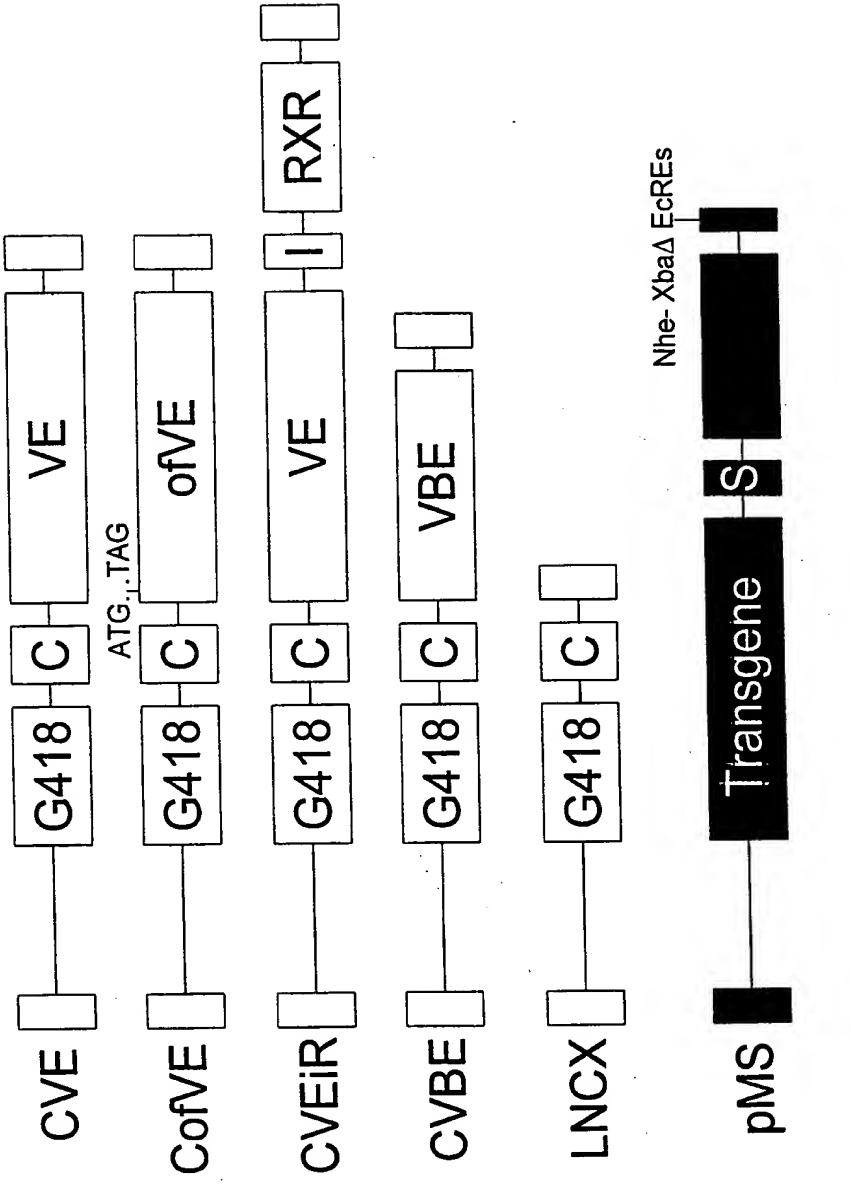

CVE
G418
C
VE
CofVE
G418
C
ofVE
ATG...TAG
CVEiR
G418
C
VE
I
RXR
CVBE
G418
C
VBE
LNCX
G418
C
pMS
Transgene
S
Nhe- XbaΔ EcREs

Figure 9A

Figure 9B-1

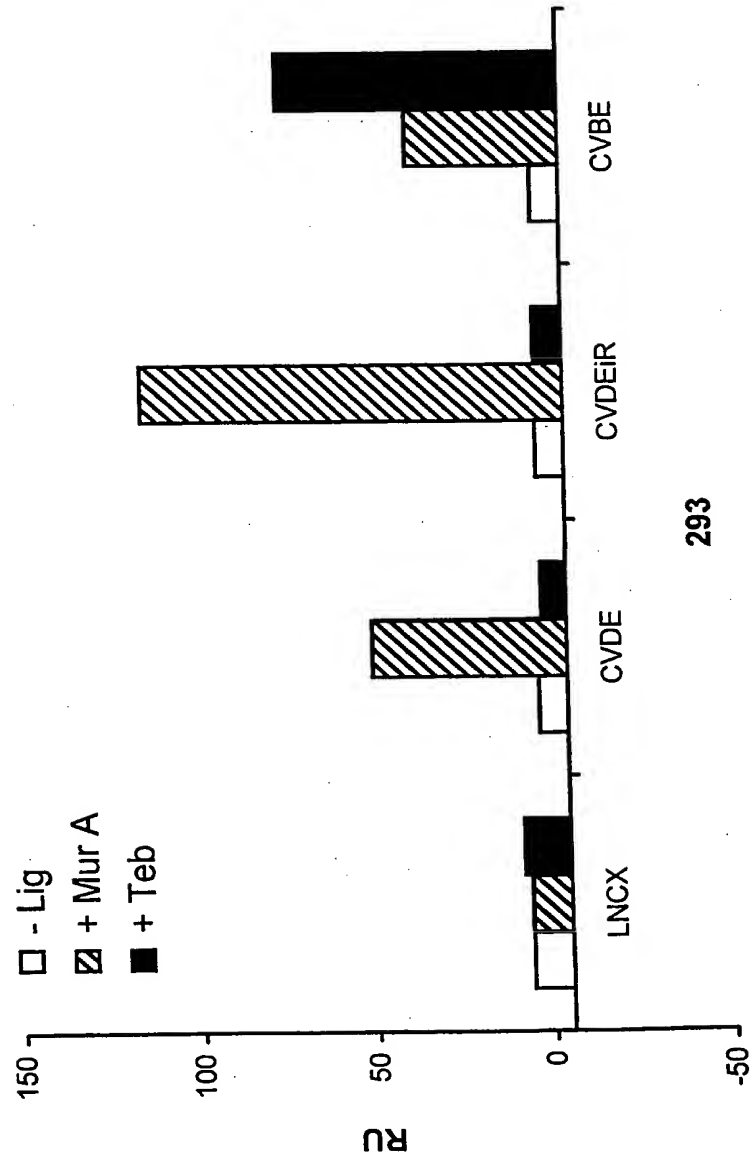


Figure 9B-2

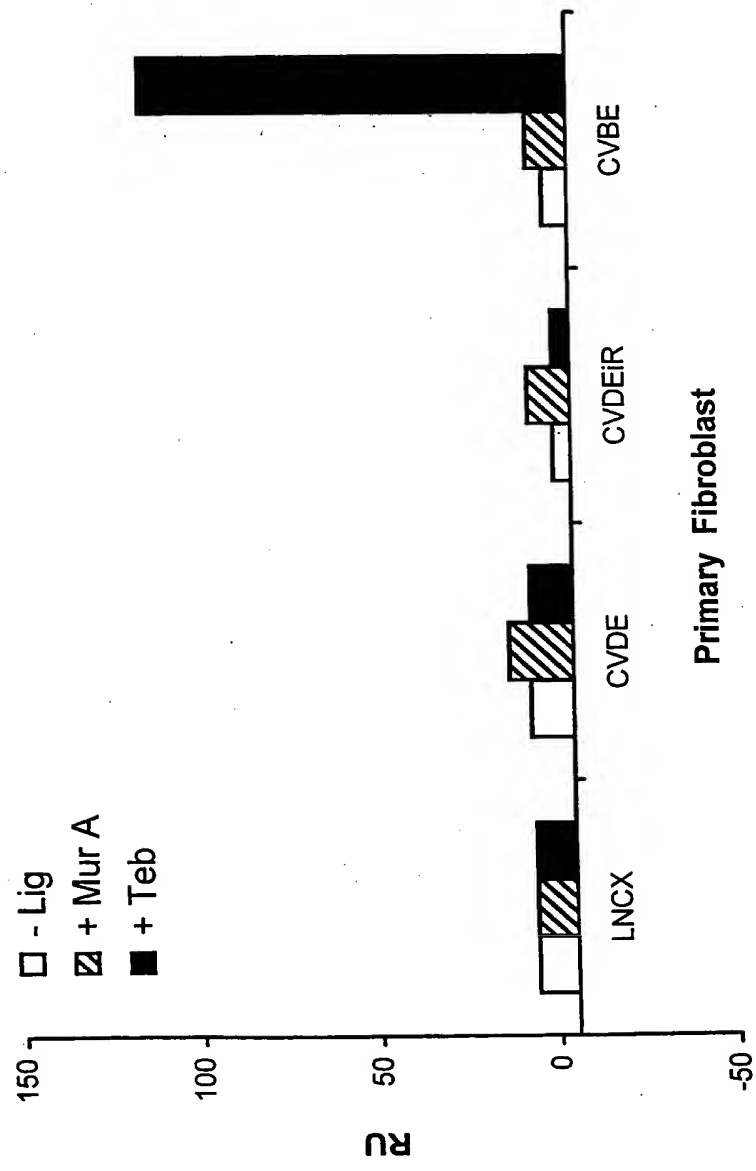

- Lig
+ Mur A
+ Teb
RU
150
100
50
0
-50
LNCX
CVDE
CVDEiR
CVBE
Primary Fibroblast

Figure 9B-3

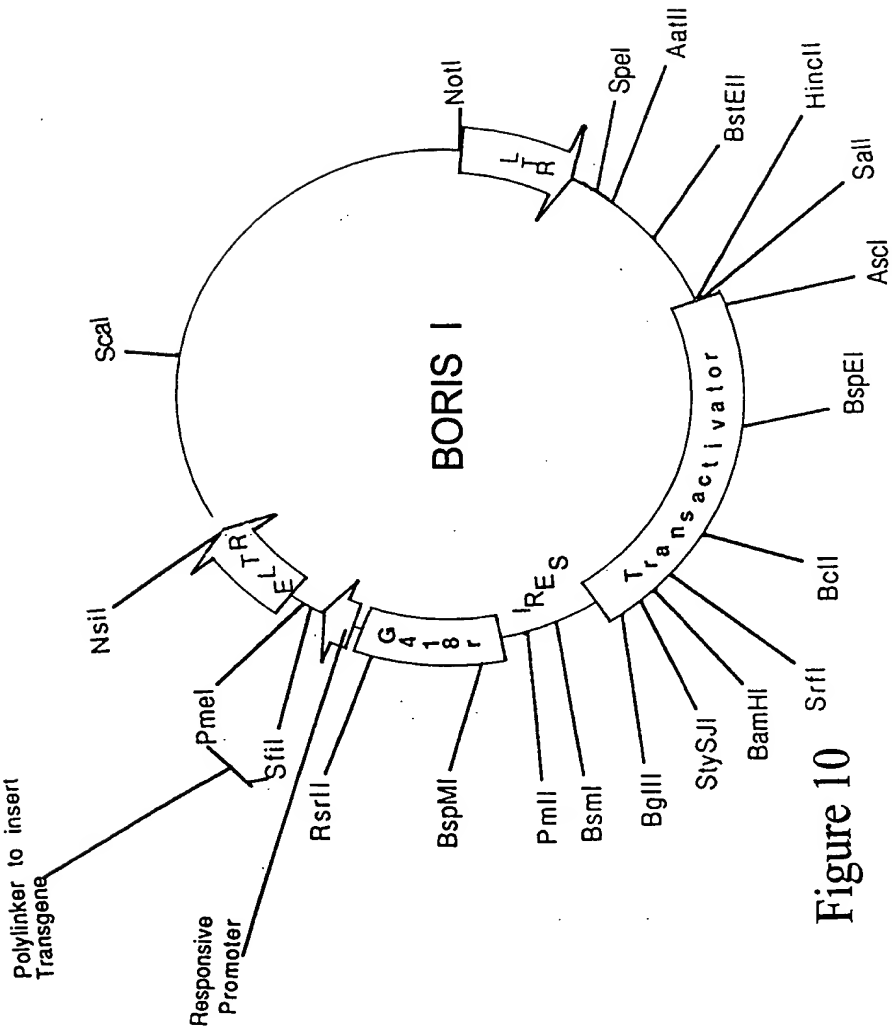

BORIS I
NotI
SpeI
AatII
BstEII
HincII
SalI
AscI
BspEI
BclI
SrfI
BamHI
StySJI
BglII
BsmI
PmlI
BspMI
RsrII
SfiI
PmeI
NsiI
ScaI
Polylinker to insert Transgene
Responsive Promoter
Transactivator
IRES
G418r
LTR
LTR

Figure 10

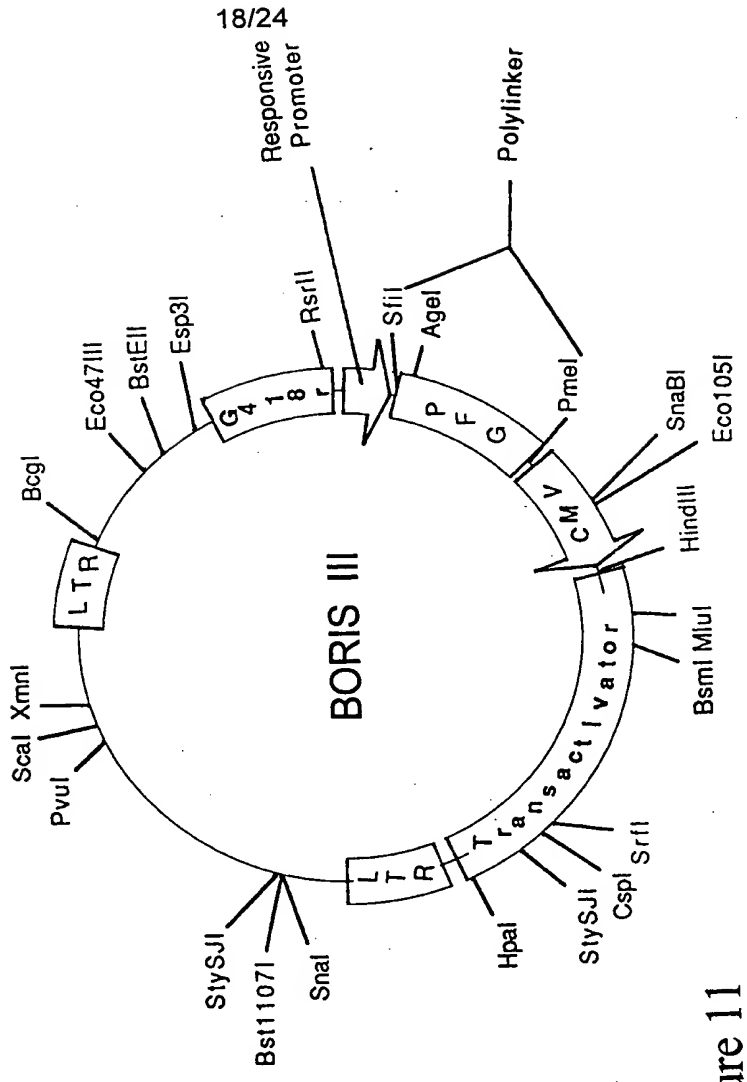

BORIS III
Responsive Promoter
Polylinker
RsrII
SfiI
AgeI
PmeI
SnaBI
Eco105I
HindIII
BsmI MluI
SrfI
CspI
StySJI
HpaI
SnaI
Bst1107I
StySJI
PvuI
ScaI XmnI
BcgI
Eco47III
BstEII
Esp3I
LTR
LTR
G418r
PFG
CMV
Transactivator

Figure 11

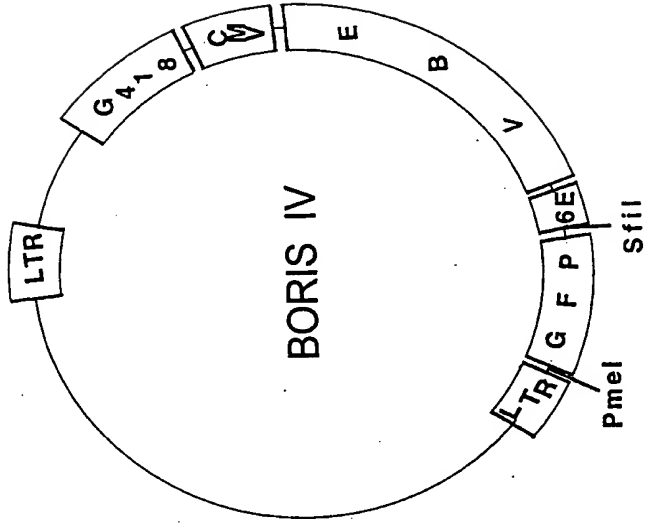


Figure 12

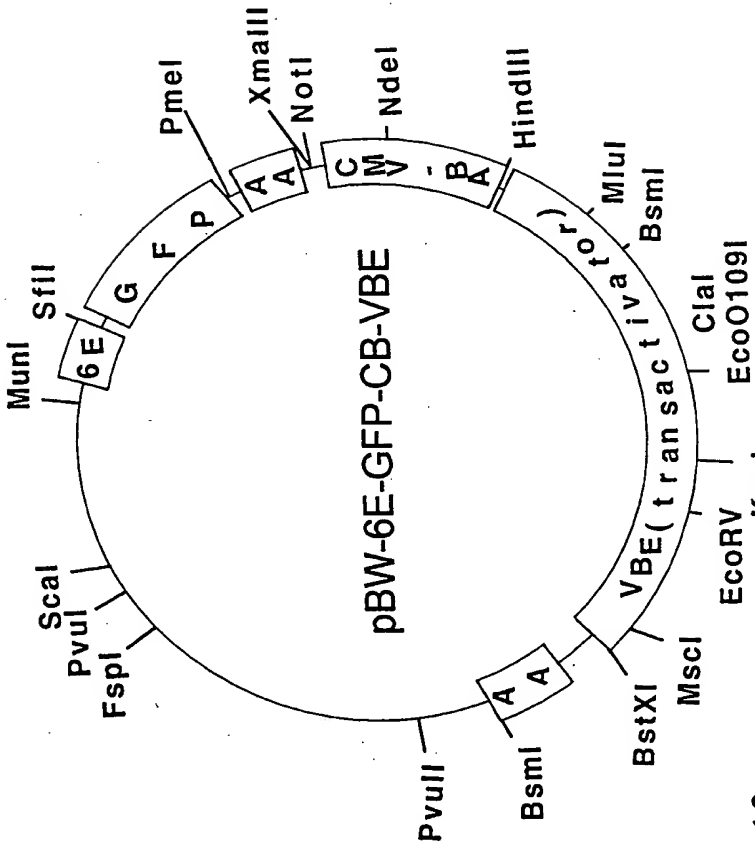

pBW-6E-GFP-CB-VBE
XmaIII
NotI
NdeI
HindIII
MluI
BsmI
ClaI
EcoO109I
KpnI
EcoRV
MscI
BstXI
BsmI
PvuII
FspI
PvuI
ScaI
MunI
SfiI
PmeI
A
A
C
M
V
-
B
A
6E
G
F
P
A
A
VBE(transactivator)

Figure 13

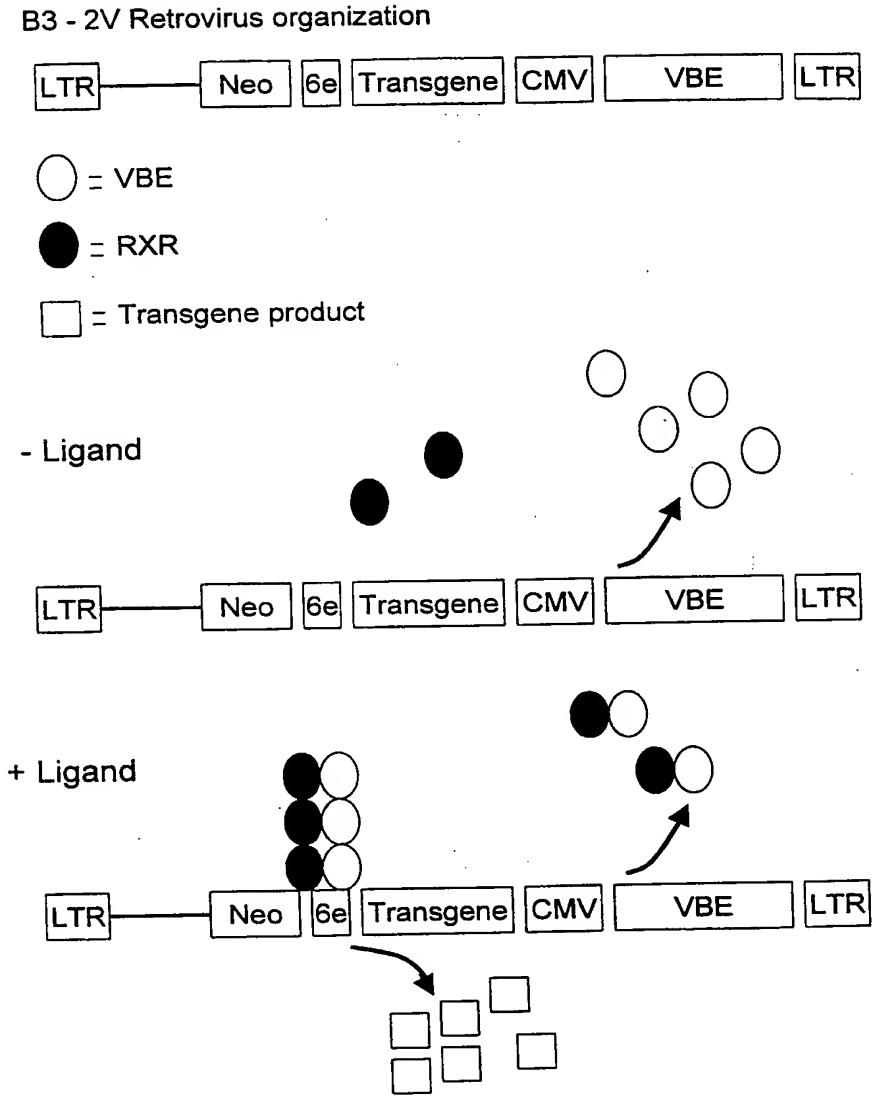
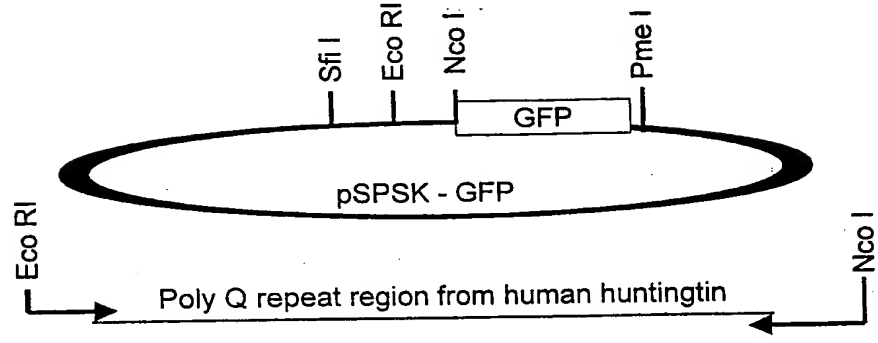

B3 - 2V Retrovirus organization
LTR
Neo
6e
Transgene
CMV
VBE
LTR
= VBE
= RXR
= Transgene product
- Ligand
+ Ligand

Figure 14

Poly - Q expression constructs

polyQ - GFP



Digest PCR product EcoRI/NcoI and insert into SPRK - GFP in - frame with the GFP ORF

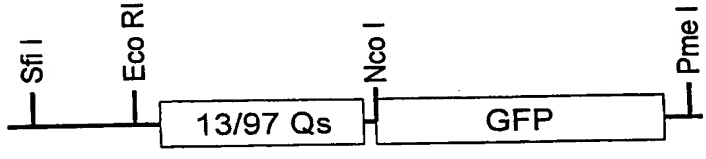


Figure 15A

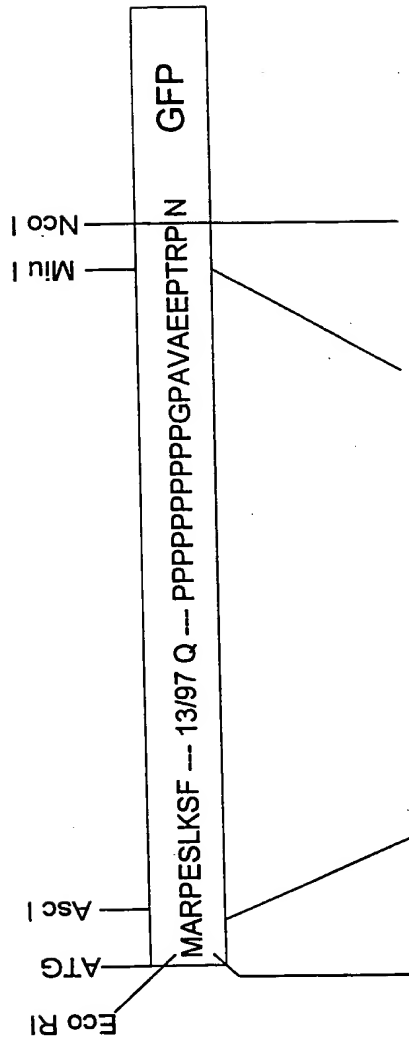

Eco RI
ATG
Asc I
MARPESLKSF — 13/97 Q — PPPPPPPPPGPAVAEEPTRP
N
GFP
Mlu I
Nco I
Met - SV40 NLS
Met - HisTag
SV40 NLS or His - Tag can be fused in frame to the N - terminus of generated contructs.
SV40 NLS
SV40 NLS or can be fused in frame down-stream of the polyQ repeat region.
All constructs and modifications are digested with Sfi I and Pme I for insertion into retroviral vectors

Figure 15B

B3 - V polyQ - GFP constructs to date

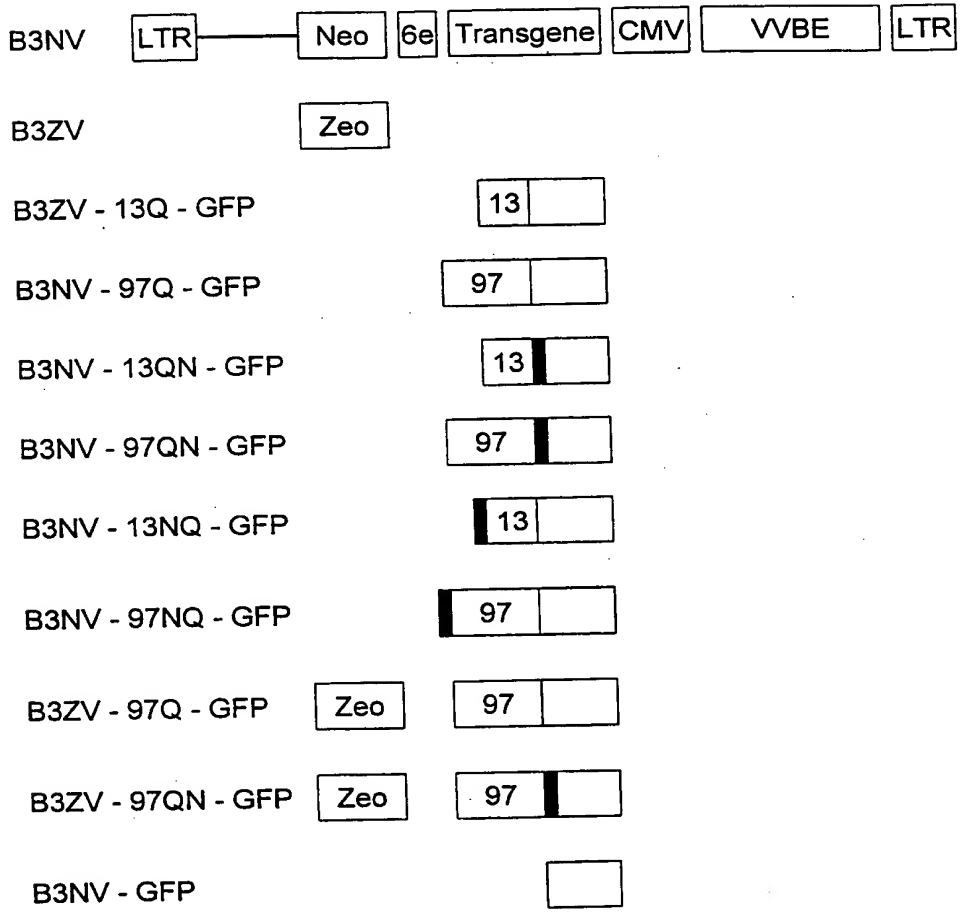


Figure 16

SEQUENCE LISTING

<110> The Salk Institute for Biological Studies

<120> Modified Lepidopteran Receptors and
Hybrid Multi-Functional Proteins for Use in Transcription
and Transgene Expression Regulation

<130> SALK2070WO

<140> PCT/US98/14215
<141> 1998-07-10

<150> 60/091874
<151> 1998-07-07

<150> 08/891,298
<151> 1997-07-10

<160> 8

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 71
<212> PRT
<213> Artificial Sequence

<220>
<223> DNA binding domain of member of steroid/thyroid
hormone receptor superfamily

<400> 1

```
Cys Xaa Xaa Cys Xaa Xaa Asp Xaa Ala Xaa Gly Xaa Tyr Xaa Xaa Xaa
 1               5                  10                  15
Xaa Cys Xaa Xaa Cys Lys Xaa Phe Phe Xaa Arg Xaa Xaa Xaa Xaa Xaa
            20                  25                  30
Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
        35                  40                  45
Xaa Xaa Xaa Lys Xaa Xaa Arg Xaa Xaa Cys Xaa Xaa Cys Arg Xaa Xaa
    50                  55                  60
Lys Cys Xaa Xaa Xaa Gly Met
65                  70
```

<210> 2
<211> 2711
<212> DNA
<213> Bombyx mori

<220>
<221> CDS
<222> (185)...(2002)

<400> 2

```
ctccgttgac gacggtcgcg cgtgcggtac gtccgttttt acggctcaaa cgtacacggt    60
aacctccgtc tctgcatcat cggcggaact cgtgaaattc gcgtgctttt ctcacctgtt   120
gaacgagttg tgttgtgact gaaaaacatc atcacaaata tcaagcttca aaactaataa   180
gtga atg aga gtc gag aac gtg gat aac gta tcg ttt gct ttg aac gga     229
     Met Arg Val Glu Asn Val Asp Asn Val Ser Phe Ala Leu Asn Gly
```

```
     1               5                  10                  15

cgc gct gac gag tgg tgt atg tct gta gag acg cgt tta gat agt tta      277
Arg Ala Asp Glu Trp Cys Met Ser Val Glu Thr Arg Leu Asp Ser Leu
                20                  25                  30

gtg cga gaa aaa agt gaa gtg aaa gcc tac gtc gga gga tgt ccc tcg      325
Val Arg Glu Lys Ser Glu Val Lys Ala Tyr Val Gly Gly Cys Pro Ser
            35                  40                  45

gta atc acg gat gct gga gcg tat gac gcg ctc ttc gac atg aga cgc      373
Val Ile Thr Asp Ala Gly Ala Tyr Asp Ala Leu Phe Asp Met Arg Arg
        50                  55                  60

cgc tgg tct aat aac ggt ggc ttc ccg ctg cga atg ctt gaa gag agc      421
Arg Trp Ser Asn Asn Gly Gly Phe Pro Leu Arg Met Leu Glu Glu Ser
    65                  70                  75

tct tca gaa gtg aca tcg tct tcg gca ctg ggt ttg cca ccg gcc atg      469
Ser Ser Glu Val Thr Ser Ser Ser Ala Leu Gly Leu Pro Pro Ala Met
80                  85                  90                  95

gtt atg tcg ccg gaa tcc ttg gcg tcg ccc gag tat cga gcc ctc gag      517
Val Met Ser Pro Glu Ser Leu Ala Ser Pro Glu Tyr Arg Ala Leu Glu
                100                 105                 110

cta tgg agc tac gat gac gga atc act tat aat aca gcc cag tct ctg      565
Leu Trp Ser Tyr Asp Asp Gly Ile Thr Tyr Asn Thr Ala Gln Ser Leu
            115                 120                 125

ctg ggt gca tgc aat atg caa cag caa cag cta caa cct cag caa cca      613
Leu Gly Ala Cys Asn Met Gln Gln Gln Gln Leu Gln Pro Gln Gln Pro
        130                 135                 140

cat cca gca cca ccg acg ctc ccc acg atg cct tta cca atg cct ccc      661
His Pro Ala Pro Pro Thr Leu Pro Thr Met Pro Leu Pro Met Pro Pro
    145                 150                 155

aca aca ccg aaa tca gaa aat gaa tcg atg tca tca ggt cga gag gaa      709
Thr Thr Pro Lys Ser Glu Asn Glu Ser Met Ser Ser Gly Arg Glu Glu
160                 165                 170                 175

ctt tcg ccg gct tca agc ata aat ggc tgc agt gct gat gct gac gcc      757
Leu Ser Pro Ala Ser Ser Ile Asn Gly Cys Ser Ala Asp Ala Asp Ala
                180                 185                 190

aga cgg cag aag aaa ggt cct gca cct cga cag caa gag gag cta tgt      805
Arg Arg Gln Lys Lys Gly Pro Ala Pro Arg Gln Gln Glu Glu Leu Cys
            195                 200                 205

ctt gtc tgc ggc gac aga gcc tcc gga tac cac tac aac gca ctg acg      853
Leu Val Cys Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr
        210                 215                 220

tgt gaa gga tgc aaa gga ttc ttc agg cgg agt gtc acc aaa aac gca      901
Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Val Thr Lys Asn Ala
    225                 230                 235

gta tat att tgt aaa ttt gga cat gcc tgt gaa atg gat atg tac atg      949
Val Tyr Ile Cys Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr Met
240                 245                 250                 255
```

```
agg agg aaa tgt caa gag tgt cga tta aag aaa tgt cta gcg gta gga      997
Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val Gly
                260                 265                 270

atg agg cct gaa tgt gtc ata cag gag ccc agt aaa aat aaa gac agg     1045
Met Arg Pro Glu Cys Val Ile Gln Glu Pro Ser Lys Asn Lys Asp Arg
            275                 280                 285

caa aga caa aag aaa gac aaa gga ata tta tta cct gtt agt acg acc     1093
Gln Arg Gln Lys Lys Asp Lys Gly Ile Leu Leu Pro Val Ser Thr Thr
        290                 295                 300

aca gtc gaa gac cac atg ccc ccg atc atg caa tgt gat cca cct ccg     1141
Thr Val Glu Asp His Met Pro Pro Ile Met Gln Cys Asp Pro Pro Pro
    305                 310                 315

ccc gag gcc gcc agg att cac gaa gtc gtc ccg agg tat ctt tcg gag     1189
Pro Glu Ala Ala Arg Ile His Glu Val Val Pro Arg Tyr Leu Ser Glu
320                 325                 330                 335

aag ctg atg gag cag aac agg cag aag aac ata cca cca ttg tcg gcg     1237
Lys Leu Met Glu Gln Asn Arg Gln Lys Asn Ile Pro Pro Leu Ser Ala
                340                 345                 350

aat cag aag tct ctg atc gcg agg ctc gtg tgg tac cag gag gga tat     1285
Asn Gln Lys Ser Leu Ile Ala Arg Leu Val Trp Tyr Gln Glu Gly Tyr
            355                 360                 365

gag cag ccc tcc gac gag gat ctc aaa aga gta acg cag act tgg cag     1333
Glu Gln Pro Ser Asp Glu Asp Leu Lys Arg Val Thr Gln Thr Trp Gln
        370                 375                 380

tcg gat gaa gag gac gag gaa tcc gat cta ccc ttc cgc cag atc acg     1381
Ser Asp Glu Glu Asp Glu Glu Ser Asp Leu Pro Phe Arg Gln Ile Thr
    385                 390                 395

gag atg acg atc tta acg gtc cag ttg atc gtc gag ttc gcc aag ggt     1429
Glu Met Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly
400                 405                 410                 415

cta ccg ggc ttt tcg aag ata tca cag tct gat caa atc acc tta tta     1477
Leu Pro Gly Phe Ser Lys Ile Ser Gln Ser Asp Gln Ile Thr Leu Leu
                420                 425                 430

aaa gcc tcg tcc agc gag gtg atg atg ctg cgg gtg gcg agg cga tac     1525
Lys Ala Ser Ser Ser Glu Val Met Met Leu Arg Val Ala Arg Arg Tyr
            435                 440                 445

gac gcc gcg tcc gac agc gtg ctg ttc gcc aac aac aag gcg tac acg     1573
Asp Ala Ala Ser Asp Ser Val Leu Phe Ala Asn Asn Lys Ala Tyr Thr
        450                 455                 460

cgc gac aac tac cgc caa ggc ggc atg gcc tac gtc atc gaa gac ctc     1621
Arg Asp Asn Tyr Arg Gln Gly Gly Met Ala Tyr Val Ile Glu Asp Leu
    465                 470                 475

cta cac ttc tgc cgg tgc atg ttc gcg atg ggc atg gac aat gtg cac     1669
Leu His Phe Cys Arg Cys Met Phe Ala Met Gly Met Asp Asn Val His
480                 485                 490                 495

ttt gca ctg ctc acg gcc atc gtt ata ttc tca gat cgg ccc ggg ctc     1717
Phe Ala Leu Leu Thr Ala Ile Val Ile Phe Ser Asp Arg Pro Gly Leu
```

```
                500                 505                 510

gag cag ccg tcg ctg gta gaa gag atc cag aga tac tac ctg aac acg      1765
Glu Gln Pro Ser Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr
            515                 520                 525

ttg cga att tac atc atc aac cag aac agc gcg tcg tcg cgc tgc gcc      1813
Leu Arg Ile Tyr Ile Ile Asn Gln Asn Ser Ala Ser Ser Arg Cys Ala
        530                 535                 540

gtg atc tac ggc agg atc ctg agc gtg ctg acc gag cta cgc acg ctc      1861
Val Ile Tyr Gly Arg Ile Leu Ser Val Leu Thr Glu Leu Arg Thr Leu
    545                 550                 555

ggc acg caa aac tcc aac atg tgc atc tcg ctg aag ctg aag aac agg      1909
Gly Thr Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys Asn Arg
560                 565                 570                 575

aag ctg ccg ccg ttc ctc gag gag atc tgg gac gtg gcg gag gtg gcc      1957
Lys Leu Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala Glu Val Ala
                580                 585                 590

acg acg cat ccc acg gtg ctg ccg ccc acc aac ccg gtg gtg cta          2002
Thr Thr His Pro Thr Val Leu Pro Pro Thr Asn Pro Val Val Leu
            595                 600                 605

tagcctccgc ccgccccagg agagaacgct catagactgg ctagttttag tgaacgtgcg   2062
ctgatccgta ttcggtgaca gattagtgat tatatgtgtt gttgaacgtt tggagagtat   2122
atatatagtg ttgacggcga ggcccgtccg gccccgtact tgtttcgttt ctgaccggat   2182
gctgcgtcgg tcgcgccctt gcgaccacga taagactact ttctataagt acgtctctaa   2242
attgaggccc caagaaccgg aagcctaact aagatacgac ggtttatttt atcacagagg   2302
aaactgaagt aattaatata ttaacagtta acacaggttc aaagcagtta ggccgctccg   2362
gacttctcat ggtctcaata aggcggacgt ataagagtta catagagaat aaaataataa   2422
tatatgaaga gatgtttcta ttggatggaa tgcgtgatgt aaagttgata gttatttta    2482
tttaccaaat taatgaagcg tcgggtgtag acctttttgta tgtgatgtgg cgaggagtgg  2542
atcgcagtgt cggccgcgtg ctctcaccaa aagcgtgcgg tcgacgctaa tagtgcgatg   2602
gttttggaat atgtttgttt atatatagtt tatgtgtgag gtgttatcgt gtcccgtcaa   2662
ttttaatttc gattcgcgtt cattcgtcct gtgctcgcta ctcagattt               2711
```

<210> 3
<211> 606
<212> PRT
<213> Bombyx mori

<400> 3

```
Met Arg Val Glu Asn Val Asp Asn Val Ser Phe Ala Leu Asn Gly Arg
 1               5                  10                  15
Ala Asp Glu Trp Cys Met Ser Val Glu Thr Arg Leu Asp Ser Leu Val
            20                  25                  30
Arg Glu Lys Ser Glu Val Lys Ala Tyr Val Gly Gly Cys Pro Ser Val
        35                  40                  45
Ile Thr Asp Ala Gly Ala Tyr Asp Ala Leu Phe Asp Met Arg Arg Arg
    50                  55                  60
Trp Ser Asn Asn Gly Gly Phe Pro Leu Arg Met Leu Glu Glu Ser Ser
65                  70                  75                  80
Ser Glu Val Thr Ser Ser Ser Ala Leu Gly Leu Pro Pro Ala Met Val
                85                  90                  95
Met Ser Pro Glu Ser Leu Ala Ser Pro Glu Tyr Arg Ala Leu Glu Leu
            100                 105                 110
Trp Ser Tyr Asp Asp Gly Ile Thr Tyr Asn Thr Ala Gln Ser Leu Leu
        115                 120                 125
Gly Ala Cys Asn Met Gln Gln Gln Gln Leu Gln Pro Gln Gln Pro His
```

```
    130                 135                 140
Pro Ala Pro Pro Thr Leu Pro Thr Met Pro Leu Pro Met Pro Pro Thr
145                 150                 155                 160
Thr Pro Lys Ser Glu Asn Glu Ser Met Ser Ser Gly Arg Glu Glu Leu
                165                 170                 175
Ser Pro Ala Ser Ser Ile Asn Gly Cys Ser Ala Asp Ala Asp Ala Arg
            180                 185                 190
Arg Gln Lys Lys Gly Pro Ala Pro Arg Gln Gln Glu Glu Leu Cys Leu
        195                 200                 205
Val Cys Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys
    210                 215                 220
Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Val Thr Lys Asn Ala Val
225                 230                 235                 240
Tyr Ile Cys Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr Met Arg
                245                 250                 255
Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val Gly Met
            260                 265                 270
Arg Pro Glu Cys Val Ile Gln Glu Pro Ser Lys Asn Lys Asp Arg Gln
        275                 280                 285
Arg Gln Lys Lys Asp Lys Gly Ile Leu Leu Pro Val Ser Thr Thr Thr
    290                 295                 300
Val Glu Asp His Met Pro Pro Ile Met Gln Cys Asp Pro Pro Pro Pro
305                 310                 315                 320
Glu Ala Ala Arg Ile His Glu Val Val Pro Arg Tyr Leu Ser Glu Lys
                325                 330                 335
Leu Met Glu Gln Asn Arg Gln Lys Asn Ile Pro Pro Leu Ser Ala Asn
            340                 345                 350
Gln Lys Ser Leu Ile Ala Arg Leu Val Trp Tyr Gln Glu Gly Tyr Glu
        355                 360                 365
Gln Pro Ser Asp Glu Asp Leu Lys Arg Val Thr Gln Thr Trp Gln Ser
    370                 375                 380
Asp Glu Glu Asp Glu Glu Ser Asp Leu Pro Phe Arg Gln Ile Thr Glu
385                 390                 395                 400
Met Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala Lys Gly Leu
                405                 410                 415
Pro Gly Phe Ser Lys Ile Ser Gln Ser Asp Gln Ile Thr Leu Leu Lys
            420                 425                 430
Ala Ser Ser Ser Glu Val Met Met Leu Arg Val Ala Arg Arg Tyr Asp
        435                 440                 445
Ala Ala Ser Asp Ser Val Leu Phe Ala Asn Asn Lys Ala Tyr Thr Arg
    450                 455                 460
Asp Asn Tyr Arg Gln Gly Gly Met Ala Tyr Val Ile Glu Asp Leu Leu
465                 470                 475                 480
His Phe Cys Arg Cys Met Phe Ala Met Gly Met Asp Asn Val His Phe
                485                 490                 495
Ala Leu Leu Thr Ala Ile Val Ile Phe Ser Asp Arg Pro Gly Leu Glu
            500                 505                 510
Gln Pro Ser Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu Asn Thr Leu
        515                 520                 525
Arg Ile Tyr Ile Ile Asn Gln Asn Ser Ala Ser Ser Arg Cys Ala Val
    530                 535                 540
Ile Tyr Gly Arg Ile Leu Ser Val Leu Thr Glu Leu Arg Thr Leu Gly
545                 550                 555                 560
Thr Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys Asn Arg Lys
                565                 570                 575
Leu Pro Pro Phe Leu Glu Glu Ile Trp Asp Val Ala Glu Val Ala Thr
            580                 585                 590
Thr His Pro Thr Val Leu Pro Pro Thr Asn Pro Val Val Leu
        595                 600                 605
```

<210> 4
<211> 17

<212> DNA
<213> Artificial Sequence

<220>
<223> GAL4 response element

<400> 4 17
cggaggactg tcctccg

<210> 5
<211> 30
<212> DNA
<213> Artificial Sequence

<400> 5 30
gagagaagct tatggcgcgc ccgaccgatg

<210> 6
<211> 29
<212> DNA
<213> Artificial Sequence

<400> 6 29
cacacacgcg tgtactcgtc aattccaag

<210> 7
<211> 1977
<212> DNA
<213> Artificial Sequence

<400> 7
atggcgcgcc cgaccgatgt cagcctgggg gacgagctcc acttagacgg cgaggacgtg 60
gcgatggcgc atgccgacgc gctagacgat ttcgatctgg acatgttggg ggacggggat 120
tccccgggtc cgggatttac cccccacgac tccgcccccc acggcgctct ggatatggcc 180
gacttcgagt ttgagcagat gtttaccgat gcccttggaa ttgacgagta cacgcgttta 240
gatagtttag tgcgagaaaa aagtgaagtg aaagcctacg tcggaggatg tccctcggta 300
atcacggatg ctggagcgta tgacgcgctc ttcgacatga gacgccgctg gtctaataac 360
ggtggcttcc cgctgcgaat gcttgaagag agctcttcag aagtgacatc gtcttcggca 420
ctgggtttgc caccggccat ggttatgtcg ccggaatcct tggcgtcgcc cgagtatcga 480
gccctcgagc tatggagcta cgatgacgga atcacttata atacagccca gtctctgctg 540
ggtgcatgca atatgcaaca gcaacagcta caacctcagc aaccacatcc agcaccaccg 600
acgctcccca cgatgccttt accaatgcct cccacaacac cgaaatcaga aaatgaatcg 660
atgtcatcag gtcgagagga actttcgccg gcttcaagca taaatggctg cagtgctgat 720
gctgacgcca gacggcagaa gaaaggtcct gcacctcgac agcaagagga gctatgtctt 780
gtctgcggcg acagagcctc cggataccac tacaacgcac tgacgtgtga aggatgcaaa 840
ggattcttca ggcggagtgt caccaaaaac gcagtatata tttgtaaatt tggacatgcc 900
tgtgaaatgg atatgtacat gaggaggaaa tgtcaagagt gtcgattaaa gaaatgtcta 960
gcggtaggaa tgaggcctga atgtgtcata caggagccca gtaaaaataa agacaggcaa 1020
agacaaaaga aagacaaagg aatattatta cctgttagta cgaccacagt cgaagaccac 1080
atgcccccga tcatgcaatg tgatccacct ccgcccgagg ccgccaggat tcacgaagtc 1140
gtcccgaggt atctttcgga gaagctgatg gagcagaaca ggcagaagaa cataccacca 1200
ttgtcggcga atcagaagtc tctgatcgcg aggctcgtgt ggtaccagga gggatatgag 1260
cagccctccg acgaggatct caaaagagta acgcagactt ggcagtcgga tgaagaggac 1320
gaggaatccg atctaccctt ccgccagatc acggagatga cgatcttaac ggtccagttg 1380
atcgtcgagt tcgccaaggg tctaccgggc ttttcgaaga tatcacagtc tgatcaaatc 1440
accttattaa aagcctcgtc cagcgaggtg atgatgctgc gggtggcgag gcgatacgac 1500
gccgcgtccg acagcgtgct gttcgccaac aacaaggcgt acacgcgcga caactaccgc 1560
caaggcggca tggcctacgt catcgaagac ctcctacact tctgccggtg catgttcgcg 1620
atgggcatgg acaatgtgca ctttgcactg ctcacggcca tcgttatatt ctcagatcgg 1680
cccgggctcg agcagccgtc gctggtagaa gagatccaga gatactacct gaacacgttg 1740
cgaatttaca tcatcaacca gaacagcgcg tcgtcgcgct gcgccgtgat ctacggcagg 1800
atcctgagcg tgctgaccga gctacgcacg ctcggcacgc aaaactccaa catgtgcatc 1860

tcgctgaagc tgaagaacag gaagctgccg ccgttcctcg aggagatctg ggacgtggcg 1920
gaggtggcca cgacgcatcc cacggtgctg ccgcccacca acccggtggt gctatag 1977

<210> 8
<211> 9072
<212> DNA
<213> Artificial Sequence

<400> 8
tatggaggga tacccgtggc acttttcggg gaaatgtgcg cggaacccct atttgtttat 60
ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc 120
aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct 180
tttttgcggc attttgcctt cctgtttttg ctcacccaga aacgctggtg aaagtaaaag 240
atgctgaaga tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta 300
agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagttc 360
tgctatgtgg cgcggtatta tcccgtattg acgccgggca agagcaactc ggtcgccgca 420
tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg 480
atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg 540
ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca 600
tgggggatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa gccataccaa 660
acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgcgc aaactattaa 720
ctggcgaact acttactcta gcttcccggc aacaattaat agactggatg gaggcggata 780
aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat 840
ctggagccgg tgagcgtggg tctcgcggta tcattgcagc actggggcca gatggtaagc 900
cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata 960
gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt 1020
actcatatat actttagatt gatttaaaac ttcatttta atttaaaagg atctaggtga 1080
agatcctttt tgataatctc atgaccaaaa tcccttaacg tgagttttcg ttccactgag 1140
cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tcctttttt ctgcgcgtaa 1200
tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt ggtttgtttg ccggatcaag 1260
agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg 1320
tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat 1380
acctcgctct gctaatcctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta 1440
ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg 1500
gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc 1560
gtgagctatg agaaagcgcc acgcttcccg aagggagaaa ggcggacagg tatccggtaa 1620
gcggcagggt cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggtatc 1680
tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg tgatgctcgt 1740
caggggggcg gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttcctggcct 1800
tttgctggcc ttttgctcac atgttctttc ctgcgttatc ccctgattct gtggataacc 1860
gtattaccgc ctttgagtga gctgataccg ctcgccgcag ccgaacgacc gagcgcagcg 1920
agtcagtgag cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt 1980
ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg ggcagtgagc 2040
gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta cactttatgc 2100
ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaaacagct 2160
atgaccatga ttacgccaag cgcgcaatta accctcacta aagggaacaa aagctggagc 2220
tagcggccgc aataaaataa aagattttat ttagtctcca gaaaaagggg ggaatgaaag 2280
accccacctg taggtttggc aagctagctt aagtaacgcc attttgcaag gcatggaaaa 2340
atacataact gagaatagag aagttcagat caaggtcagg aacagatgga acagctgaat 2400
atgggccaaa caggatatct gtggtaagca gttcctgccc cggctcaggg ccaagaacag 2460
atggaacagc tgaatatggg ccaaacagga tatctgtggt aagcagttcc tgccccggct 2520
cagggccaag aacagatggt ccccagatgc ggtccagccc tcagcagttt ctagaacaag 2580
ggttcaatgc acttgtcgct agaacaaggg ttcaatgcac ttgtcgctag aacaagggtt 2640
caatgcactt gtcgctagaa caagggttca atgcacttgt cgagaaccat cagatgtttc 2700
cagggtgccc caaggacctg aaatgaccct gtgccttatt tgaactaacc aatcagttcg 2760
cttctcgctt ctgttcgcgc gcttctgctc cccgagctca ataaaagagc ccacaacccc 2820
tcactcgggg cgccagtcct ccgattgact gagacgcgtc ggggtacccg tattcccaat 2880
aaagcctctt gctgttgcat ccgaatcgtg gactcgctga tccttgggag ggtctcctca 2940
gattgattga ctgcccacct cgggggtctt tcatttggag gttccaccga gattggagac 3000
ccctgcctag ggaccaccga ccccccccgc cgggaggtaa gctggccagc ggtcggtttc 3060
gtgtctgtct ctgtctttgt gcgtgtttgt gccggcatct aatgtttgcg cctgcgtctg 3120
tactagttag ctaactagct ctgtatctgg cggacccgtg gtggaactga cgagttctga 3180

acacccggcc gcaaccctgg gagacgtccc agggactttg ggggccgttt ttgtggcccg 3240
acctgaggaa gggagtcgat gtggaatccg accccgtcag gatatgtggt tctggtagga 3300
gacgagaacc taaaacagtt cccgcctccg tctgaatttt tgctttcggt ttggaaccga 3360
agccgcgcgt cttgtctgct gcagcatcgt tctgtgttgt ctctgtctga ctgtgtttct 3420
gtatttgtct gaaaattagg gccagactgt taccactccc ttaagtttga ccttaggtca 3480
ctggaaagat gtcgagcgga tcgctcacaa ccagtcggta gatgtcaaga agagacgttg 3540
ggttaccttc tgctctgcag aatggccaac ctttaacgtc ggatggccgc gagacggcac 3600
ctttaaccga gacctcatca cccaggttaa gatcaaggtc ttttcacctg gcccgcatgg 3660
acacccagac caggtcccct acatcgtgac ctgggaagcc ttggcttttg acccccctcc 3720
ctgggtcaag ccctttgtac accctaagcc tccgcctcct cttcctccat ccgccccgtc 3780
tctccccctt gaacctcctc gttcgacccc gcctcgatcc tccctttatc cagccctcac 3840
tccttctcta ggcgccggaa ttcggcgcgt gtggcctcga acaccgagcg accctgcagc 3900
gacccgctta acagcgtaag cttgtcgacg gtatcgataa gcttatggcg cgcccgaccg 3960
atgtcagcct gggggacgag ctccacttag acggcgagga cgtggcgatg gcgcatgccg 4020
acgcgctaga cgatttcgat ctggacatgt tgggggacgg ggattccccg ggtccgggat 4080
ttacccccca cgactccgcc ccctacggcg ctctggatat ggccgacttc gagtttgagc 4140
agatgtttac cgatgccctt ggaattgacg agtacacgcg tttagatagt ttagtgcgag 4200
aaaaaagtga agtgaaagcc tacgtcggag gatgtccctc ggtaatcacg gatgctggag 4260
cgtatgacgc gctcttcgac atgagacgcc gctggtctaa taacggtggc ttcccgctgc 4320
gaatgcttga agagagctct tcagaagtga catcgtcttc ggcactgggt ttgccaccgg 4380
ccatggttat gtcgccggaa tccttggcgt cgcccgagta tcgagccctc gagctatgga 4440
gctacgatga cggaatcact tataatacag cccagtctct gctgggtgca tgcaatatgc 4500
aacagcaaca gctacaacct cagcaaccac atccagcacc accgacgctc cccacgatgc 4560
ctttaccaat gcctcccaca acaccgaaat cagaaaatga atcgatgtca tcaggtcgag 4620
aggaactttc gccggcttca agcataaatg gctgcagtgc tgatgctgac gccagacggc 4680
agaagaaagg tcctgcacct cgacagcaag aggagctatg tcttgtctgc ggcgacagag 4740
cctccggata ccactacaac gcactgacgt gtgaaggatg caaaggattc ttcaggcgga 4800
gtgtcaccaa aaacgcagta tatatttgta aatttggaca tgcctgtgaa atggatatgt 4860
acatgaggag gaaatgtcaa gagtgtcgat taaagaaatg tctagcggta ggaatgaggc 4920
ctgaatgtgt catacaggag cccagtaaaa ataaagacag gcaaagacaa aagaaagaca 4980
aaggaatatt attacctgtt agtacgacca cagtcgaaga ccacatgccc ccgatcatgc 5040
aatgtgatcc acctccgccc gaggccgcca ggattcacga agtcgtcccg aggtatcttt 5100
cggagaagct gatggagcag aacaggcaga agaacatacc accattgtcg gcgaatcaga 5160
agtctctgat cgcgaggctc gtgtggtacc aggagggata tgagcagccc tccgacgagg 5220
atctcaaaag agtaacgcag acttggcagt cggatgaaga ggacgaggaa tccgatctac 5280
ccttccgcca gatcacggag atgacgatct taacggtcca gttgatcgtc gagttcgcca 5340
agggtctacc gggcttttcg aagatatcac agtctgatca aatcacctta ttaaaagcct 5400
cgtccagcga ggtgatgatg ctgcgggtgg cgaggcgata cgacgccgcg tccgacagcg 5460
tgctgttcgc caacaacaag gcgtacacgc gcgacaacta ccgccaaggc ggcatggcct 5520
acgtcatcga agacctccta cacttctgcc ggtgcatgtt cgcgatgggc atggacaatg 5580
tgcactttgc actgctcacg gccatcgtta tattctcaga tcggcccggg ctcgagcagc 5640
cgtcgctggt agaagagatc cagagatact acctgaacac gttgcgaatt tacatcatca 5700
accagaacag cgcgtcgtcg cgctgcgccg tgatctacgg caggatcctg agcgtgctga 5760
ccgagctacg cacgctcggc acgcaaaact ccaacatgtg catctcgctg aagctgaaga 5820
acaggaagct gccgccgttc ctcgaggaga tctgggacgt ggcggaggtg gccacgacgc 5880
atcccacggt gctgccgccc accaacccgg tggtgctata gcctccgccc gccccaggag 5940
agaacgctca tagactggct agttttagtg aacgtgcgct gatccgtatt cggtgacaga 6000
ttagtgatta tatgtgttgt tgaacgtttg gagagtatat atatagtgtt atcgaattcc 6060
gcccctctcc ctcccccccc cctaacgtta ctggccgaag ccgcttggaa taaggccggt 6120
gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggcaat gtgagggccc 6180
ggaaacctgg ccctgtcttc ttgacgagca ttcctagggg tcttttcccct ctcgccaaag 6240
gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct tcttgaagac 6300
aaacaacgtc tgtagcgacc ctttgcaggc agcggaaccc cccacctggc gacaggtgcc 6360
tctgcggcca aaagccacgt gtataagata cacctgcaaa ggcggcacaa ccccagtgcc 6420
acgttgtgag ttggatagtt gtggaaagag tcaaatggct ctcctcaagc gtattcaaca 6480
aggggctgaa ggatgcccag aaggtacccc attgtatggg atctgatctg gggcctcggt 6540
gcacatgctt tacgtgtgtt tagtcgaggt taaaaaacgt ctaggccccc cgaaccacgg 6600
ggacgtggtt ttcctttgaa aaacacgatg ataatatggc cacaaccatg gcatgattga 6660
acaagatgga ttgcacgcag gttctccggc cgcttgggtg gagaggctat tcggctatga 6720
ctgggcacaa cagacaatcg gctgctctga tgccgccgtg ttccggctgt cagcgcaggg 6780
gcgcccggtt ctttttgtca agaccgacct gtccggtgcc ctgaatgaac tgcaggacga 6840
ggcagcgcgg ctatcgtggc tggccacgac gggcgttcct tgcgcagctg tgctcgacgt 6900

tgtcactgaa gcgggaaggg actggctgct attgggcgaa gtgccggggc aggatctcct 6960
gtcatctcac cttgctcctg ccgagaaagt atccatcatg gctgatgcaa tgcggcggct 7020
gcatacgctt gatccggcta cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg 7080
agcacgtact cggatggaag ccggtcttgt cgatcaggat gatctggacg aagagcatca 7140
ggggctcgcg ccagccgaac tgttcgccag gctcaaggcg cgcatgcccg acggcgagga 7200
tctcgtcgtg acccatggcg atgcctgctt gccgaatatc atggtggaaa atggccgctt 7260
ttctggattc atcgactgtg gccggctggg tgtggcggac cgctatcagg acatagcgtt 7320
ggctaccgt gatattgctg aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct 7380
ttacggtatc gccgctcccg attcgcagcg catcgccttc tatcgccttc ttgacgagtt 7440
cttctgagcg ggactctggg gttcgaacct gtaggtttgg caagctagaa caagggttca 7500
atgcacttgt cgctagaaca agggttcaat gcacttgtcg ctagaacaag ggttcaatgc 7560
acttgtcgct agaacaaggg ttcaatgcac ttgtcgagaa ccatgggtcg agtaggcgtg 7620
tacggtggga ggcctatata agcagagctc gtttagtgaa ccgtcagatc gcctggagac 7680
gccatccaac gctgttttgc ctccaaaaga caccgggacc gaggccgcct gggcccgttt 7740
aaacatcgat aaaataaaag attttattta gtctccagaa aaaggggga atgaaagacc 7800
ccacctgtag gtttggcaag ctagcttaag taacgccatt ttgcaaggca tggaaaaata 7860
cataactgag aatagagaag ttcagatcaa ggtcaggaac agatggaaca gctgaatatg 7920
ggccaaacag gatatctgtg gtaagcagtt cctgccccgg ctcagggcca agaacagatg 7980
gaacagctga atatgggcca aacaggatat ctgtggtaag cagttcctgc cccggctcag 8040
ggccaagaac agatggtccc cagatgcggt ccagccctca gcagtttcta gaacaagggt 8100
tcaatgcact tgtcgctaga acaagggttc aatgcacttg tcgctagaac aagggttcaa 8160
tgcacttgtc gctagaacaa gggttcaatg cacttgtcga gaaccatcag atgtttccag 8220
ggtgccccaa ggacctgaaa tgaccctgtg ccttatttga actaaccaat cagttcgctt 8280
ctcgcttctg ttcgcgcgct tctgctcccc gagctcaata aaagagccca caacccctca 8340
ctcgggcgc cagtcctccg attgactgag tcgccgggt acccgtattc ccaataaacc 8400
tcttgctgat gcatggtacc caattcgccc tatagtgagt cgtattacgc gcgctcactg 8460
gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttacccaact taatcgcctt 8520
gcagcacatc ccccttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct 8580
tcccaacagt tgcgcagcct gaatggcgaa tggaaattgt aagcgttaat attttgttaa 8640
aattcgcgtt aaatttttgt taaatcagct catttttaa ccaataggcc gaaatcggca 8760
aaatcccctta taaatcaaaa gaatagaccg agatagggtt gagtgttgtt ccagtttgga 8820
acaagagtcc actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc 8880
agggcgatgg cccactacgt gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc 8940
gtaaagcact aaatcggaac cctaaaggga gcccccgatt tagagcttga cggggaaagc 9000
cggcgaacgt ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg 9060
caagtgtagc ggtcacgctg cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac 9072
agggcgcgtc ag

INTERNATIONAL SEARCH REPORT

Inter. onal Application No
PCT/US 98/14215

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C12N15/86 C12N5/10 A01K67/027 C07K14/72
A61K48/00 G01N33/566

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SWEVERS L ET AL.: "The silkmoth homolog of the Drosophila ecdysone receptor (B1 isoform): Cloning and analysis of ecxpression during follicular cell differentiation" INSECT BIOCHEMISTRY AND MOLECULAR BIOLOGY, vol. 25, no. 7, July 1995, pages 857-866, XP002086543 ISSN 0965-1748 see figures 1 and 2 --- -/--	1-17, 21-25, 28-30, 34-38, 40,41, 47-69

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
2 December 1998	21/12/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Cupido, M

1

INTERNATIONAL SEARCH REPORT

Inter onal Application No
PCT/US 98/14215

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 37609 A (ZENECA LIMITED) 28 November 1996 see the whole document	1-3, 5-17, 21-25, 28-30, 34-38, 40,41, 47-69
P,X	SUHR T S ET AL.: "High level transactivation by a modified Bombyx ecdysone receptor in mammalian cells without exogenous retinoid X receptor" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., vol. 95, no. 14, 7 July 1998, pages 7999-8004, XP002086544 WASHINGTON US see the whole document	1-17, 21-30, 34-69, 73-76, 82-84

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 98/ 14215

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 47-76 and claims 78 and 79 insofar as in vivo methods are concerned are directed to a method of treatment of the human/ animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

information on patent family members

Inten ...nal Application No
PCT/US 98/14215

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9637609 A	28-11-1996	AU 5771696 A	11-12-1996
		CA 2219121 A	28-11-1996
		CN 1191568 A	26-08-1998
		CZ 9703722 A	18-03-1998
		EP 0828829 A	18-03-1998
		NO 975419 A	22-01-1998
		PL 323587 A	14-04-1998